VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIZERTAČNÍ PRÁCE

Brno, 2020

RNDr. BEDŘICH SMETANA

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV MATEMATIKY
DEPARTMENT OF MATHEMATICS

ALGEBRAIZACE A PARAMETRIZACE PŘECHODOVÝCH RELACÍ MEZI STRUKTUROVANÝMI OBJEKTY S APLIKACEMI V OBLASTI NEURONOVÝCH SÍTÍ
ALGEBRAIZATION AND PARAMETERIZATION TRANSITION RELATIONS BETWEEN STRUCTURED OBJECTS WITH APPLICATIONS IN THE FIELD OF NEURAL NETWORKS

DIZERTAČNÍ PRÁCE
DOCTORAL THESIS

AUTOR PRÁCE
AUTHOR
RNDr. Bedřich Smetana

VEDOUCÍ PRÁCE
ADVISOR
prof. RNDr. Jan Chvalina, DrSc.

BRNO 2020

ABSTRAKT

V disertační práci je studováno modelování činnosti neuronové sítě se zaměřením na vícevrstvou dopřednou neuronovou síť (MLP – Multi Layer Perceptron). V této často užívané struktuře neuronové sítě je nově využito časově proměnných neuronů (time-varying neurons) spolu s analogií při modelování hyperstruktur lineárních diferenciálních operátorů. Za pomocí koncového lematu a definované hyperoperace je definována u dané přechodové funkce hyperstruktura složená z neuronů. U těchto struktur zkoumáme jejich vlastnosti s akcentem na uspořádané struktury.

KLÍČOVÁ SLOVA

Struktura a funkce dopředné neuronové sítě, Matematický model neuronové sítě,Časově proměnné neurony, Struktury lineárních diferenciálních operátorů, Analogie struktur časově proměnných neuronů, Teorie hyperstruktur a automatů, Modelovací funkce, Spojnicový prostor.

ABSTRACT

The dissertation thesis investigates the modeling of the neural network activity with a focus on a multilayer forward neural network (MLP – Multi Layer Perceptron). In this often used structure of neural networks, time-varying neurons are used, along with an analogy in modeling hyperstructures of linear differential operators. Using a finite lemma and defined hyperoperation, a hyperstructure composed of neurons is defined for a given transient function. There are examined their properties with an emphasis on structures with a layout.

KEYWORDS

Structure and function of feed-forward neural network, Mathematical model of neural network, Time-varying neurons, Structures of linear differential operators, Analogy of structures of time-varying neurons, Theory of hyperstructures and automata, Modeling functions, Join space.

SMETANA, Bedřich. *Algebraizace a parametrizace přechodo- vých relací mezi strukturovanými objekty s aplikacemi v oblasti neuronových sítí*. Brno, 2020, 93 s. Dizertační práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav matematiky. Vedoucí práce: prof. RNDr. Jan Chvalina, DrSc.

PROHLÁŠENÍ

Prohlašuji, že svou dizertační práci na téma „Algebraizace a parametrizace přechodových relací mezi strukturovanými objekty s aplikacemi v oblasti neuronových sítí“ jsem vypracoval samostatně pod vedením školitele dizertační práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené dizertační práce dále prohlašuji, že v souvislosti s vytvořením této dizertační práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno ....

podpis autora

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu dizertační práce panu prof. RNDr. Janu Chvalinovi, DrSc., za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Obsah

Seznam obrázků

Úvod

Práce je motivována ideou zobecnění funkce časově proměnného neuronu a pomocí analogie s lineárními diferenciálními operátory a jejich algebraickými strukturami tvořit a zkoumat algebraické struktury těchto neuronů na podkladě funkce dopředné neuronové sítě - vícevrstvého perceptronu MLP (vícevrstvý perceptron – multilayer perceptron). Již profesor Otakar Borůvka [17] a zejména přední představitel této matematické školy profesor František Neuman [91, 92, 93] ve svých publikacích z oblasti lineárních diferenciálních rovnic využíval algebraizace a geometrizace studovaných objektů s akcentem na pologrupy a grupy. Také fenomén neuronových sítí a jejich činnosti a nalezení nových souvislostí s algebraickými strukturami a rovněž možnost nového pohledu na tyto sítě byla další motivací této práce, zejména poznatky Tima Koskely [76] z oblasti časově proměnných neuronů, Geoffreye Hintona [110] a Christophera M. Bishopa [16] z oblasti funkcionality vícevrstvých neuronových sítí.

Hlavním cílem této práce je konstrukce modelu umělého neuronu v dopředné vícevrstvé neuronové síti s vybranou přenosovou funkcí a s využitím poznatků z oblasti lineárních diferenciálních operátorů vytvořit hyperstuktury umělých neuronů a tím vytvořit potenciál k novému pohledu na funkcionalitu neuronových sítí zvolené struktury a přenosové funkce.

V průběhu zpracování projektu této práce byly nejprve vybrány a popsány algebraické struktury použité v dalším výzkumu umělých neuronů [65]. Mnoho zajímavých výsledků týkajících se zkoumání slučitelných vztahů na algebrách bylo získáno I. Chajdou, J. Niederlem, B. Pondělíčkem, J. Rachůnkem, B. Zelinkou a dalšími (např. [104, 45, 46, 47, 48]). V článcích výše uvedených autorů je substituční vlastnost (SP – substituční vlastnost – substitution property) binárních vztahů na dané algebře považována za typ kompatibility. Pro výzkum chování umělých neuronů vzhledem k přenosové funkci je významný pojem funkční parametrizace nebo funkční rozklad. V teorii obecných vstupně-výstupních systémů, kterými jsou triády S = (X, S, Y), jsou pojmy jako funkční rozklad obecných systémů a globální reakce, definovány jako důležité koncepty. V této souvislosti jsou zkoumány kompatibilní relace na vytvořených strukturách a hyperstrukturách umělých neuronů včetně parametrizace pomocí endomorfismů. V kontextu funkční struktury dopředné neuronové sítě byly zkoumány možnosti konceptů kompatibility binárních relací topologických prostorů a jejich možného uplatnění v hyperstrukturách. Další poznatky ve zkoumaných relacích se zabývají již zmíněnou substituční vlastností.

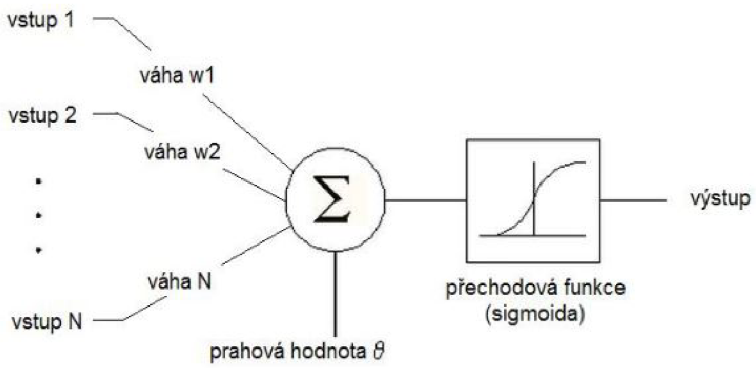
Stěžejní pro další výzkum algebraických struktur tvořených umělými neurony dopředné neuronové sítě bylo využití vlastností zvolené přenosové funkce spolu s analogií s konceptem lineárních diferenciálních operátorů. Získané poznatky umožnily vytvořit koncept grup a hypergrup lineárních diferenciálních operátorů a jejich vlastností. Nejprve z vlastností propagační funkce a Cliffordova neuronu (CN) společně s poznatky Hintona [108, 110] byl vytvořen funkční zápis činnosti neuronové jednotky. Dalším krokem byla volba vstupů a vah neuronové sítě jako funkce argumentu t z lineárně uspořádané časové množiny T. Tato matematická konstrukce a zobecnění neuronové sítě umožnila další výzkum vlastností algebraických struktur umělých neuronů.

1 Základní poznatky – přehled současného stavu

1.1 Problematika vývoje neuronových sítí

Výzkum neuronových sítí je disciplina, která mimo matematiky a statistiky spojuje již od počátku další vědní obory, od psychologie a fyziologie, přes fyzikální aspekty šíření a filtraci signálů a také elektromagnetické vlastnosti látek. Vývoj v této oblasti nebyl vždy přímočarý a vzhledem k rozsáhlé a dynamicky se vyvíjející vědní disciplině s mnoha odlišně fungujícími příklady sítí a také vzhledem k vlastnímu fungování sítí je nezbytné popsat v základních rysech mimo vývoj také důležité zástupce neuronových sítí.

Moderní pohled na neuronové sítě započal v roce 1940 s prací Warrena McCullocha a Waltera Pittse, kteří ukázali, že sítě z umělých neuronů mohou v zásadě počítat s libovolnou matematickou nebo logickou funkcí. Vytvořili přitom velmi jednoduchý matematický model neuronu, který je základní buňkou nervového systému [10, 20, 110, 85]. Číselné hodnoty parametru v tomto modelu byly bipolární, tj. z množiny $\{-1, 0, 1\}$. V roce 1957 Frank Rosenblatt navázal na tyto poznatky tzv. perceptronem, který je zobecněním McCullochova a Pittsova modelu neuronu pro reálný číselný obor parametrů. Tato síť byla postavena Frankem Rosenblattem a jeho týmem [107] s cílem naučit ji schopnost rozpoznávat vzory. Pro tento model navrhl učící algoritmus, o kterém matematicky dokázal, že pro daná tréninková data nalezne po konečném počtu kroků odpovídající váhový vektor parametrů (pokud existuje) nezávisle na jeho počátečním nastavení. Činnost této sítě analogicky uplatňuje mimo jiné také myšlenky akademika Viléma Laufbergera a jeho vzruchové teorie fungování paměti [77].



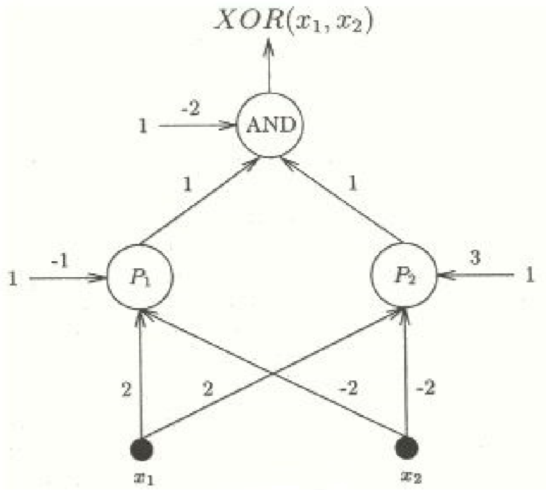
Obr. 1.1: McCullochův a Pittsův model neuronu [86]

Původní jednoduchá struktura formálního neuronu s binárním výstupem a jednovrstevná neuronová síť – jednovrstvý perceptron – nevyhovovala velkému množství úloh, které nelze vyjádřit jednoduchým lineárním vektorovým prostorem. Pokračováním vývoje v tomto směru je vícevrstvý perceptron (MLP).

Využití zejména třívrstvého perceptronu a poznatky použití perceptronu při rozpoznávání vzorů se snahou analyzovat různé koncepty včetně omezení popisuje Frank Rosenblatt v knize o neurovýpočtech Principles of Neurodynamics [106] a bývá proto považován za zakladatele oboru neuronových sítí.

Následovala práce Donalda Hebba [36], který navrhl, na základě podmíněného reflexu (teorie Pavlova), mechanismus pro učení formálních neuronů [36]. Krátce po objevu perceptronu Bernard Widrow se svými studenty vyvinul další typ neuronového výpočetního prvku, který nazval ADALINE (adaptivní lineární neuron – ADAptive LInear NEuron) [120].

Po období prudkého vývoje a přehnaných očekávání došlo vlivem práce Marvina Minského a Seymoura Paperta [84] z roku 1969 k útlumu prostředků, podporujících výzkum neuronových sítí. Tito vědci využili svého vlivu, aby diskreditovali výzkum neuronových sítí, nacházející se v krizi, ve snaze převést finanční zdroje z této oblasti na jiný výzkum v oboru umělé inteligence. Základem kampaně byla jejich výzkumná zpráva publikovaná v roce 1969 pod názvem Perceptrons. V této knize Minsky a Papert využili pro svoji argumentaci známého triviálního faktu, že jeden perceptron nemůže počítat jednoduchou logickou funkci XOR.

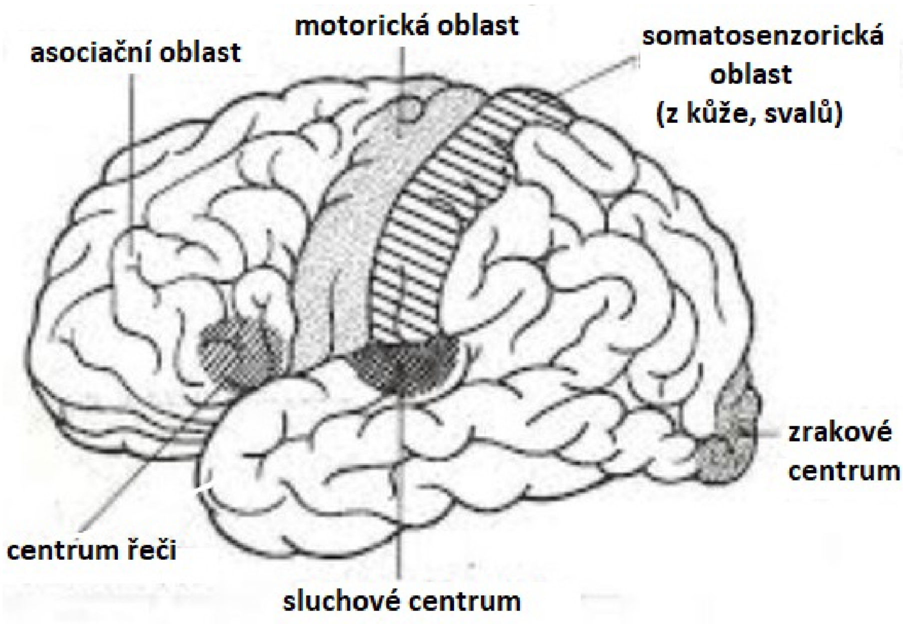


Obr. 1.2: Funkce XOR - dvouvrstvý perceptron [122]

Tato funkce, označovaná jako exkluzivní logický součet, poskytuje pro dva vstupy logickou výstupní hodnotu 1 jen tehdy, když se oba vstupy liší. Řešením problému je využití dvouvrstvé sítě se třemi neurony, pro tuto síť nebyl znám učící algoritmus, který podle autorů vzhledem ke složitosti nebylo možné vytvořit, což bylo všeobecně přejato a považováno za matematicky dokázané. Další vývoj proto probíhal skrytě například v oblasti adaptivních algoritmů. Až na počátku 80. let badatelé v oblasti neurovýpočtů začali podávat vlastní grantové projekty zaměřené na vývoj neuropočítačů a jejich aplikace. Podporu stejně jako například při vývoji internetu poskytla americká grantová agentura DARPA (Agentura ministerstva obrany pro pokročilé výzkumné projekty – Defense Advanced Research Projects Agency), která začala finančně podporovat výzkum zejména díky činnosti programového manažera Ira Skurnicka, brzy následovaly další organizace podporující základní i aplikovaný výzkum. Objevila se řada zajímavých talentů jako Shun-Ichi Amari, James Anderson, Kunihiko Fukushima, Stephen Grossberg, Teuvo Kohonen a David Willshaw. Velkou zásluhu na dalším vývoji oboru neuronových sítí měl světově uznávaný fyzik John Hopfield, který se v osmdesátých letech začal zabývat neurovýpočty [40]. V letech 1982 až 1984 dokázal souvislost některých modelů neuronových sítí s fyzikálními modely magnetických materiálů. Později podle něj byly nazvány tzv. Hopfieldovy sítě fungující na principu autoasociativní paměti. Síť má pevnou topologii s „n“ neurony, které jsou zapojeny cyklicky. Všechny neurony jsou zároveň výstupní. V roce 1986 publikovali své výsledky badatelé z tzv. PDP skupiny (Parallel Distributed Processing Group). V článku Davida Rumelharta, Geoffreyho Hintona a Ronalda Williamse popsali učící algoritmus zpětného šíření chyby (backpropagation) pro vícevrstvou neuronovou síť [108]. Tím padly důvody opuštění této oblasti výzkumu, obsažené v knize Minského a Paperta. Následovaly další významné objevy včetně využití alternativních neuronů. Roku 1987 se v San Diegu konala první větší konference s výhradním zaměřením na neuronové sítě. IEEE International Conference on Neural Networks uvítala přes 1700 účastníků s jasným výsledkem založení mezinárodní společnosti pro výzkum neuronový sítí INNS (International Neural Network Society). Vzniká řada světoznámých časopisů jako jsou Neural Networks, Neural Computation, IEEE Transactions on Neural Networks. I Česká republika získává v roce 1991 mezinárodní časopis Neural Networks World. Od roku 1987 řada univerzit na celém světě zakládá nové výzkumné ústavy s tímto zaměřením. Mimo vývoje neuropočítačů se v dnešní době se především výpočty simulují na klasických PC stanicích, avšak množství aplikací nalézá cestu i do mobilních zařízení. Neuronové sítě jsou oblastí zájmu výzkumu vojenských, civilních, komerčních organizací a univerzitních pracovišť s časopisy, které se soustřeďují na tuto problematiku. Ani poté však není vývoj neuronových sítí přímočarý. Koncem devadesátých let až po rok 2005 jsou neuronové sítě zatlačeny do pozadí jinými modely (support vector machines (SVM)). Od roku 2006 až po současnost nastává renesance neuronových sítí s akcentem na hluboké sítě (mnoho vrstev). Hlubokým učením (Deep Learning) označujeme soubor algoritmů využívaných pro strojové učení. Obvykle se používají u neuronových sítí. Neuronové sítě, které obsahují mnoho vrstev a využívají metod Hlubokého učení, se nazývají Deep Neural Networks. Hluboké neuronové sítě s učením bez učitele nebo s generativním učením jsou určeny k zachycení korelace pozorovaných nebo viditelných dat, pro analýzu vzoru nebo pro účely syntézy. Síť se tedy snaží nalézt spojitosti mezi daty na základě samotných dat a nikoli na základě jejich popisu nebo klasifikační třídy. Nejrozšířenější a nejběžnější jsou modely založeny na energii (Energy-based deep models). Typickým představitelem je základní model autoenkodéru (Deep autoenencoders). Dalším význačným zástupcem této kategorie je Hluboký Boltzmannův stroj (Deep Boltzmann machine). Jedná se o zvláštní případ obecného Boltzmannova stroje (Boltzmann machine). Tento model obsahuje mnoho vrstev skrytých proměnných. Proměnné v rámci jedné vrstvy nejsou vzájemně propojeny. Každá následující vrstva Hlubokého Boltzmannova stroje zachycuje složitější korelace v datech, a proto mají velký potenciál naučit se složitě vnitřní reprezentace dat. Myšlenku vícevrstevných sítí publikoval ukrajinský teoretik Alexej Ivachněnko se svými kolegy v roce 1966. Jeho algoritmus Group Method of Data Handling byl vyvinut jako nástroj pro nalezení závislostí v komplexních nelineárních vícerozměrných systémech, jejich predikci (předpovědi) a aproximaci [71]. Propagátory hlubokého učení jsou Yoshua Bengio, Geoffrey Hinton a Yann Le Cun z CIFAR (Kanadský institut pokročilého výzkumu) v Torontu, známý mezi kolegy též jako „kanadská mafie".

1.2 Biologická inspirace neuronových sítí

Vztahy mezi vnějším prostředím a organismem, i mezi jeho částmi, a příslušná reakce na vnější podněty i na vnitřní stavy organismu probíhá šířením vzruchů z jednotlivých čidel, tzv. receptorů, které umožňují přijímat mechanické, tepelné, chemické a světelné podněty, směrem k jiným nervovým buňkám, které tyto signály zpracovávají a přivádí k příslušným výkonným orgánům, tzv. efektorům. Tyto vzruchy se po projekčních drahách, kde dochází k prvnímu předzpracování, kompresi a filtraci informace, dostávají až do mozkové kůry, která je nejvyšším řídicím centrem nervového systému. Na povrchu mozku můžeme rozlišit celkem šest primárních vzájemně propojených projekčních oblastí odpovídajících přibližně smyslům, ve kterých dochází k paralelnímu zpracování informace. Neuron – nervová buňka, je základní funkční a histologická jednotka nervové tkáně. Skládá se z těla – soma s jádrem, a z dostředivých (dendritů) a odstředivých (axonů) výběžků. Z axonu obvykle odbočuje řada větví, tzv. terminálů, zakončených blánou, která se převážně stýká s výběžky, tzv. trny, dendritů jiných neuronů. K přenosu informace pak slouží výše zmíněné unikátní mezineuronově rozhraní, tzv. (chemická) synapse [73, 75].



Obr. 1.3: Struktura mozku [75]

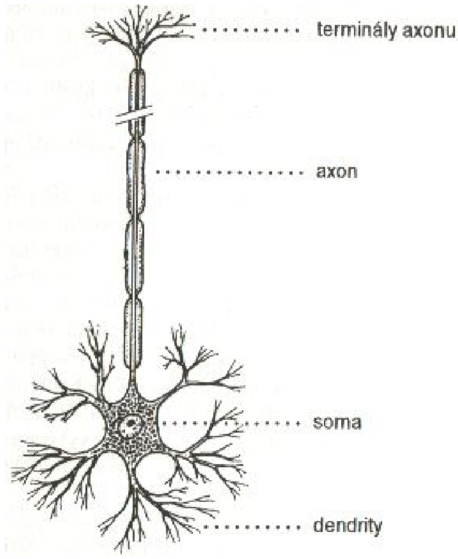
Základní myšlenkou při konstrukci neuronových sítí je analogie s činností lidského mozku. Tato oblast byla předmětem zájmu zakladatelů kybernetiky Norberta Wienera a Johna von Neumanna. Činnost lidského mozku je velmi složitá záležitost, jak ukazují i populárně publikované poznatky patologů [75], a jeho výzkum je také ovlivněn etickými hledisky. Základním stavebním kamenem nervové tkáně je specializovaná buňka-neuron, který přenáší impuls při překročení prahové hodnoty pomocí synapsí elektrochemicky. Synapse jsou kulové nebo knoflíkovité výstupky. Počet synapsí je několik tisíc až set tisíc, počet propojení nervové buňky mozku je až pět tisíc.

Na základě analogie byl vytvořen koncept formálního neuronu. Princip činnosti formálního neuronu je v literatuře objasňován na základě rozlišení dvou znaků – na podkladě percepce, rozeznávání znaků u perceptronu jako dichotomické rozdělení vstupních dat na dvě množiny představující prvky dvou částí vektorového prostoru. Obecně je funkce umělých neuronových sítí daná zobrazením:

$$\{\vec{x}\} \to \{\vec{y}\},$$

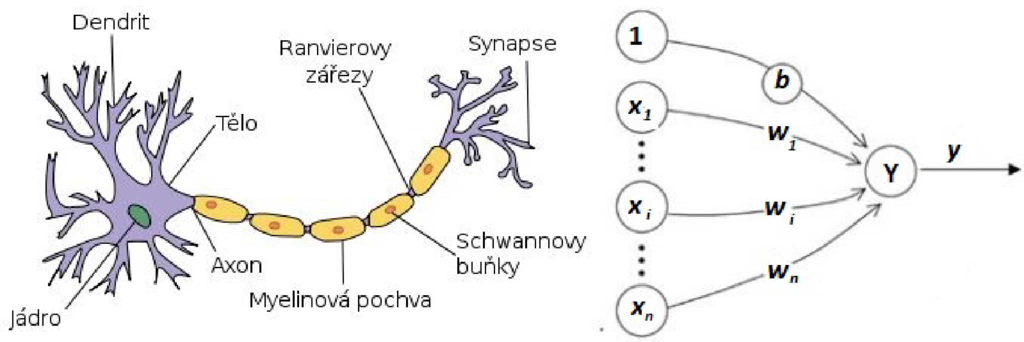
kde $\vec{x}$ je vektor vstupního signálu ze vstupního vektorového prostoru, $\vec{y}$ je vektor výstupního signálu z výstupního vektorového prostoru, které mohou mít různé dimenze. Činnost neuronové sítě lze charakterizovat:

- velkým počtem současně pracujících formálních neuronů,
- paralelním zpracováním, které kromě jiného může radikálně zvýšit rychlost zpracování,



Obr. 1.4: Biologický neuron [73]

- neurony typicky pracují zpravidla (téměř)autonomně, asynchronně s distribuovaným řízením,
- neurony jsou hustě propojeny, přičemž spoje jsou charakterizovány různými vahami, které jsou proměnné a způsobují se schopností učení flexibilitu a odolnost sítě.



Obr. 1.5: Biologický a formální neuron [114]

Základem matematického modelu neuronové sítě je tedy formální neuron. Obecně má n reálných vstupů, které modelují dendrity a určující vstupní vektor

$$\vec{x} = (x_1, ..., x_n).$$

Tyto vstupy jsou ohodnocovány reálnými synaptickými vahami, které tvoří vektor

$$\vec{w} = (w_1, ..., w_n).$$

Formální úpravou lze docílit toho, že přenosová funkce bude mít nulový práh a vlastní práh neuronu se záporným znaménkem budeme chápat jako váhu, tzv. bias $w_0 = {}^{-}\theta$ dalšího formálního vstupu $x_0 = 1$ s konstantní jednotkovou hodnotou.

Název přenosové funkce	Vstupně-výstupní relace	Ikona	Funkce v MATLAB
Skoková přenosová funkce	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		hardlim
Symetrická skoková přenosová funkce	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		hardlims
Lineární přenosová funkce	$a = n$		purelin
Saturovaná lineární přenosová funkce	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		satlin
Symetrická saturovaná lineární přenosová f.	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$		satlins
Standardní logistická sigmoida	$a = \frac{1}{1+e^{-n}}$		logsig
Hyperbolický tangens	$a = \frac{e^{n}-e^{-n}}{e^{n}+e^{-n}}$		tansig
Pozitivní lineární	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		poslin
Soutěživá	$a = 1$ neuron with max n $a = 0$ all other neurons	C	compet

Obr. 1.6: Tabulka přenosových funkcí [34]

Přiložená tabulka je výčtem používaných přenosových funkcí včetně symboliky a názvu v MATLABu, jehož moduly (Matlab Simuliuk a Neural Network Toolbox) se často využívají k simulaci neuronových sítí. Při konstrukci neuronových sítí je možné kombinovat struktury s odlišnými přenosovými funkcemi.

Označení hodnoty vnitřního potenciálu je y_in, kde

$$\mathbf{y\_in} = \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{w_i x_i}.$$

po dosažení hodnoty $\mathbf{b}(\mathbf{w_0} = \mathbf{b} - \mathbf{bias})$ vyvolává výstup (stav) $\mathbf{y}$ neuronu $\mathbf{Y}$, který modeluje elektrický puls axonu. Nárůst výstupní hodnoty $y = y\_in$ při dosažení hodnoty potenciálu $\mathbf{b}$ je dán aktivační (přenosovou) funkcí f.

Výběr aktivační funkce značnou mírou předurčuje úlohu sítě. Zápis výstupní hodnoty neuronu v maticové formě lze vyjádřit jako:

$$\mathbf{y} = \mathbf{W} * \mathbf{x} + \mathbf{b},$$

kde matice pro případ jednoho neuronu má jen jeden řádek. Výstupní a vstupní produkt umělých neuronů může být stejným způsobem interpretován jako vektory vstupních nebo výstupních lineárních vektorových prostorů. Nyní může být neuronový výstup vícevrstvé neuronové sítě psán podle podmínek zápisu maticové formy jako:

$$\mathbf{y^n} = \mathbf{f^n}(\mathbf{W^n}\mathbf{f^{n-1}}(\mathbf{W^{n-1}}\mathbf{f^{n-2}} \cdots (\mathbf{W^1}\mathbf{x} + \mathbf{b^1}) + \mathbf{b^2}) \cdots + \mathbf{b^n})$$

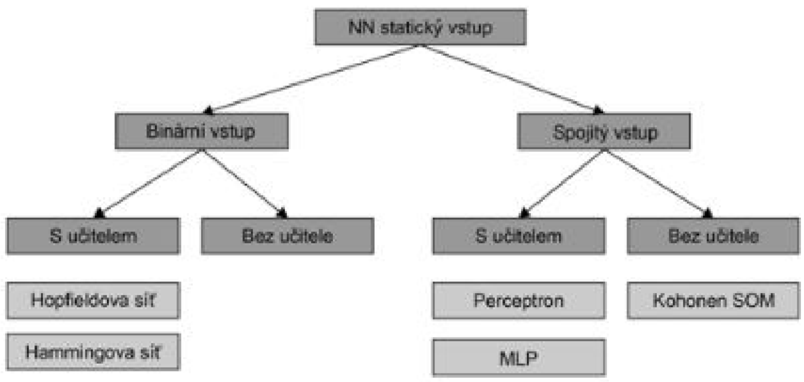
pro n-vrstvou neuronovou síť. U výstupů lze rozlišit minimálně:

- Binární výstup – využívaný zejména pro klasifikační úlohy, kdy výstupní vrstva obsahuje počet neuronů, který odpovídá počtu cílových kategorií. Síť se poté učí, aby pro každý vstupní vektor reagovala odezvou, jež bude exkluzivně obsahovat jedničku na místě pro rozpoznanou kategorii.
- Reálný výstup – je typický pro aproximační úlohy funkcí apod.

Technický prakticky realizovaný model, původně tvořený soustavou potenciometrů a motorů [107], napodoboval funkci biologického neuronu, kde přenos signálu probíhá elektrochemicky. K přenosu informace zde slouží unikátní mezineuronové rozhraní, synapse. Z funkčního hlediska lze synapse rozdělit na excitační, které umožňují rozšíření vzruchu v nervové soustavě, a na inhibiční, které způsobují jeho útlum [73, 75]. Paměťová stopa v nervové soustavě vzniká pravděpodobně zakódováním synaptických vazeb na cestě mezi receptorem (čidlem orgánu) a efektorem (výkonným orgánem). Míra podráždění neuronů prostřednictvím impulzů od jiných neuronů po dosažení určité prahové hodnoty způsobí vznik impulzu, kterým přes axon a synapse podráždí další neuron. Obdobně také matematický model může pracovat s pevnou prahovou hodnotou [106, 114]. Biologická inspirace hraje při vývoji neuronových sítí důležitou roli při analogickém modelování struktury umělé neuronové sítě a postupném výzkumu fenoménu umělé inteligence.

1.2.1 Rozdělení neuronových sítí

Ke klasifikaci neuronových sítí můžeme přistupovat z různých směrů pohledu. Používaným přístupem je dělení sítí dle formy vstupních dat, jak ukazuje diagram z článku „An introduction to computing with neural nets“ od Richarda P. Lippmanna uveřejněný v časopise IEEE ASP MAGAZINE v dubnu 1987 [79]. V tomto schematu dělení dále využíváme rozdělení na sítě s učitelem a bez učitele a posléze klasifikátory neuronové sítě na základě nejbližších klasických statistických algoritmů, charakterizovaných rozdělením náhodné veličiny, kde klasifikátor rozhoduje, která třída výsledků nejlépe, tj. s nejvyšší pravděpodobností reprezentuje vstupní data. V tomto pojetí dle Lippmanna je neuronová síť charakterizovaná čistou topologií (počet neuronů a jejich vzájemné propojení v síti určuje tzv. architekturu (topologii) neuronové sítě)[1], charakteristikami uzlů (umělých neuronů) a pravidly týkající se výcviku nebo učení. Mimo adaptačního mechanismu učení sítě je důležitá fáze předpřípravy dat a to zejména u rozpoznávání obrazů a řeči, kde nelze zajistit trénovanost sítě dostatečnou saturací vzory. V této situaci se využívá snížení počtu parametrů zavedením pomocných proměnných (features). Paměť umělého neuronu není samostatná jednotka, ale je rozprostřená ve vstupní části formou váhových koeficientů. Pomocí těchto koeficientů je systém schopný zapamatovat si informace. Toto pojetí odpovídá starší představě o funkci paměti jako zafixovanému způsobu spojení synapsí, nebo také jako synaptické síly, kdy vzruch může být opakovaně přehrán, kde hraje důležitou úlohu část mozku zvaná hipokampus. V poslední době probíhá výzkum, který ukazuje důležitost ribonukleové kyseliny (RNA) u senzorických neuronů a možnost chemického přenosu některých částí paměti, zatím prověřeno u vhodného druhu laboratorních mořských plžů Zej [9]. Vzhledem k tomu, že tato substance ovlivňuje jádro neuronu, lze zde najít analogii s využitím časově proměnných neuronů v umělých neuronových sítích.

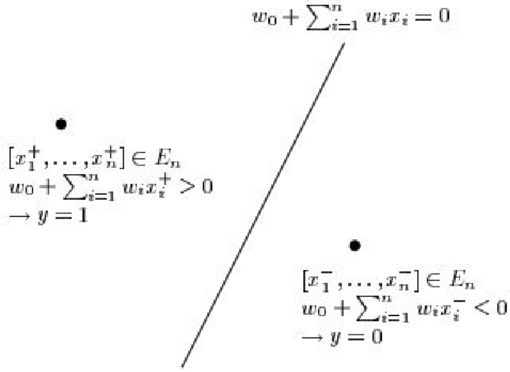


Obr. 1.7: Neuronové sítě [79]

U novějších publikací se využívá pro interpretaci neuronových sítí poznatků z oblastí lineárních vektorových prostorů a geometrické interpretace činnosti sítě. V tomto pojetí je neuronová síť nástrojem pro přenos dat ze vstupního do výstupního vektorového prostoru. Obvykle užívaným příkladem pro objasnění činnosti sítě

[1]Topologie (z řeckého topos - místo a logos - studie) je obor matematiky, opírající se o velmi obecný výklad pojmu prostor (topologický prostor). Studuje takové vlastnosti útvarů, které se nemění při oboustranně spojitých transformacích [87]. Topologie sítí se zabývá zapojením různých prvků do počítačových sítí a zachycením jejich skutečné (reálné) a logické (virtuální) podoby (datové linky, síťové uzly) [33]. Jako taková je součástí teorie grafů a zasahuje tedy i do matematiky.

je rozpoznávání dvou prvních písmen abecedy v analogii s činností mozku žáka první třídy. K lepšímu pochopení funkce jednoho neuronu v tomto pojetí nám pomůže geometrická představa načrtnutá na obrázku.



Obr. 1.8: Geometrická interpretace funkce neuronu [114]

Vstupy neuronu lze chápat jako souřadnice bodu v n-rozměrném euklidovském, tzv. vstupním prostoru E_n. V tomto prostoru má rovnice nadroviny (např. v E_2 přímka, v E_3 rovina) tvar: $w_0 + \sum_{i=1}^{n} w_i x_i = 0$. Tato nadrovina dělí vstupní prostor na dva poloprostory. Souřadnice bodů $[x_1^+, \ldots, x_n^+]$, které leží v jednom poloprostoru, splňují následující nerovnost:

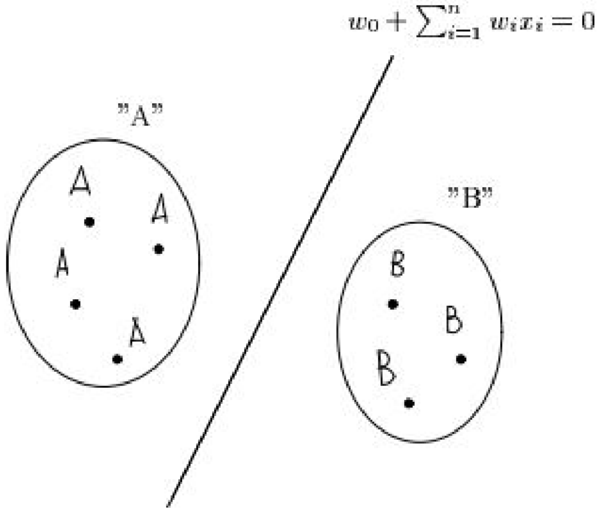
$$w_0 + \sum_{i=1}^{n} w_i x_i^+ > 0,$$

body $[x_1^-, \ldots, x_n^-]$ pak splňují nerovnost

$$w_0 + \sum_{i=1}^{n} w_i x_i^- < 0.$$

Pokud leží bod v prvním poloprostoru, je neuron aktivní, v druhém pasivní. Synaptické váhy včetně biasu jsou koeficienty nadroviny. V příkladu školáka pak rozdělení může být znázorněno takto:

V případě analogie se školákem jsou váhy nastaveny náhodně a neuron neurčí všechny hodnoty správně. Teprve při úpravě vah a natočení nadroviny dojde k úpravě (za pomoci učitele). V případě vícevrstvé sítě se skládají účinky jednotlivých neuronů. Šíření a zpracování informace na cestě v síti je umožněno změnou stavů neuronů ležících na této cestě. Stavy všech neuronů v síti určují tzv. stav neuronové sítě a synaptické váhy všech spojů představují tzv. konfiguraci neuronové sítě. Neuronová síť se v čase vyvíjí, mění se propojení a stav neuronů, adaptují se váhy. V souvislosti se změnou těchto charakteristik v čase je celková dynamika neuronové



Obr. 1.9: Neuronová síť – rozpoznání znaků [113]

sítě v této interpretaci rozdělena do tří dynamik a tří režimů práce sítě: organizační (změna topologie), aktivní (změna stavu) a adaptivní (změna konfigurace). Pro specifikaci konkrétního modelu neuronové sítě stačí, když definujeme jeho organizační, aktivní a adaptivní dynamiku.

- Organizační dynamika specifikuje architekturu sítě a její případnou změnu. Změna topologie se většinou uplatňuje v rámci adaptivního režimu tak, že síť je v případě potřeby rozšířena o další neurony a příslušné spoje. Avšak organizační dynamika převážně předpokládá pevnou architekturu neuronové sítě, která se již nemění. Z pohledu organizační dynamiky rozlišujeme v zásadě dva typy architektury: cyklická (resp. rekurentní) a acyklická (resp. dopředná) síť. V případě cyklické topologie existuje v síti skupina neuronů, která je zapojena v kruhu (tzv. cyklus).
- Aktivní dynamika specifikuje počáteční stav sítě a způsob jeho změny v čase při pevné topologii a konfiguraci. V aktivním režimu se na začátku nastaví stavy vstupních neuronů na tzv. vstup sítě a zbylé neurony jsou v uvedeném počátečním stavu. Všechny možné vstupy, resp. stavy sítě, tvoří tzv. vstupní prostor, resp. stavový prostor neuronové sítě. Po inicializaci stavu sítě probíhá vlastní výpočet. Spojitý model je zadán diferenciální rovnicí, v praxi většinou jde o diskrétní proces. Podle toho, zda neurony mění svůj stav nezávisle na sobě nebo je jejich aktualizace řízena centrálně, rozlišujeme asynchronní a synchronní modely neuronových sítí.
- Adaptivní dynamika specifikuje počáteční konfiguraci sítě a způsob změny vah v síti v čase. Všechny možné konfigurace sítě tvoří tzv. váhový prostor neuronové sítě. V adaptivním režimu jsou na začátku nastaveny váhy všech spojů

v síti na počáteční konfiguraci (např. náhodně). Po inicializaci konfigurace sítě probíhá vlastní adaptace. Podobně jako v aktivní dynamice se obecně uvažuje spojitý model se spojitým vývojem konfigurace neuronové sítě v čase, kdy váhy sítě jsou (spojitou) funkcí času, která je obvykle v adaptivní dynamice zadána diferenciální rovnicí. Většinou se však předpokládá diskrétní čas adaptace. Cílem adaptace je nalézt konfiguraci sítě ve váhovém prostoru, která by v aktivním režimu realizovala předepsanou funkci. Jestliže aktivní režim sítě se využívá k vlastnímu výpočtu funkce sítě pro daný vstup, pak adaptivní režim slouží k učení („programování“) této funkce. Existují stovky úspěšných učících algoritmů pro různé modely neuronových sítí. Nejznámější a nejpoužívanější učící algoritmus je backpropagation – algoritmus zpětného šíření chyby – pro vícevrstvou neuronovou síť.

Dalším kritériem rozdělení neuronových sítí může být způsob učení neuronové sítě, zda se učí síť sama, bez cizí pomoci, nebo zda k tomu používá vnější asistenci. Podle toho rozlišujeme:

- učení bez učitele,
- učení s učitelem.

Oba tyto druhy učení jsou používány v praxi při učení umělých neuronových sítí a neuropočítačů na nich založených.

1.2.2 Příklady neuronových sítí

Oblasti využití neuronových sítí

Pro praktické testování i provoz neuronových sítí lze použít programy s moduly pro tvorbu neuronových sítí na konvenčních počítačích. Možnosti tohoto software ukazují i možnosti použití neuronových sítí. Jedním z kvalitních a známých softwarů je STATISTICA od Stat Soft, podle kterého lze možnosti využití shrnout. Od tohoto software je možné navíc nalézt informace k tzv. automatizovaným neuronovým sítím, které představují pomocníka pro tvorbu neuronové sítě do jisté míry eliminující nutnost hlubších znalostí o konfiguraci sítě a její tvorby. Menu programu umožňuje využít neuronové sítě v těchto oblastech analýzy:

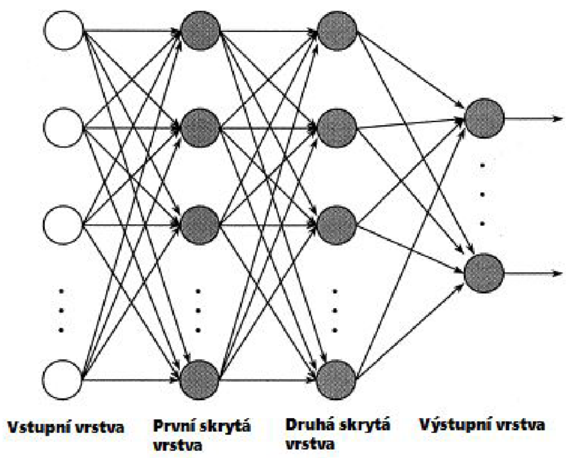
- **Regrese** – regresní analýza se zabývá předpovídáním spojité proměnné na základě vstupů (spojitých či kategorických prediktorů – nezávislých proměnných).
- **Klasifikace** – neboli zařazování do tříd. Na základě klasifikace úrovně cílové proměnné a kombinace vstupů, které ke konkrétnímu výsledku vedou, bude vytvořen model, který dokáže klasifikovat nová data. Sem lze zařadit i motivační příklad s rozpoznáváním vzorů při různých zdrojích například ručně psaných rukopisů. Dalšími častými příklady jsou bankovní problémy, které řeší otázky

bonity klienta, schopnost jeho splácení, ale také problémy typu poslat či neposlat pacienta na podrobné vyšetření, detekce spamu apod.

- **Časové řady(regrese)** – slouží k modelování spojitých proměnných, které prochází vývojem v čase, resp. časových řad. V této situaci můžeme vybrat buď pouze jedinou závislou proměnnou, kde model bude vycházet ze zpožděných hodnot této časové řady, nebo máme možnost zvolit další proměnné, které budou tuto řadu vysvětlovat.
- **Časové řady (klasifikace)** – tento typ analýzy použijeme tehdy, je-li naše cílová (závislá) proměnná kategorické povahy. Závislou proměnnou lze vysvětlovat opět pouze svým „historickým“ průběhem v čase, případně je možné zvolit další spojité i kategorické prediktory jako nezávislé vysvětlující proměnné.
- **Shluková analýza** – tento typ analýzy nepoužívá závislou proměnnou (učení bez učitele), cílem je detekovat netriviální shluky v datech. Jde o tzv. Kohonenovu síť. Vstupem jsou pouze hodnoty (vstupních) nezávislých proměnných. Tento typ problému je součástí analytické metody Data miningu a Text miningu, jako metody získávání informací buď z obecně datové nebo jen textově datové množiny dat často i z otevřených zdrojů, kde se s výhodou uplatňují neuronové sítě.

Technologii neuronových sítí lze použít prakticky v každé situaci, v níž je cílem nalezení neznámé proměnné nebo vlastnosti na základě známých pozorování nebo naměřených hodnot (tzn. v nejrůznějších typech regresí, klasifikací a časových řad), je-li k dispozici dostatečné množství historických dat a existují-li mezi nimi objektivní vztahy nebo množina vztahů (neuronové sítě jsou relativně tolerantní k šumovým jevům) existuje také reálná šance, že je bude možno prostřednictvím neuronových sítí modelovat. Je možné zmínit celou řadu oblastí, jako je automatické řízení, kybernetika, umělá inteligence, optimalizace a expertní systémy, řízení procesů, kde lze nalézt příklady jako je monitorování strojního zařízení v průmyslových procesech a průběžné nastavování řídicích parametrů, samoobslužné mechanismy subsystémů např. řízení skleníků, budov, dopravní signalizace, samoučící se zahradní a zemědělská technika, rozpoznávání poruch strojů, následné plánování preventivní údržby, provozování strojů (např. odhady spotřeby paliv na základě snímaných měření), plánování využití prostoru, strategie maximalizace výkrmu hospodářských zvířat, rozpoznávání kvality výrobku. Dalšími rozsáhlými oblastmi využití jsou zpracování obrazu a řeči s příklady – optické rozpoznávání textů, písma, obrazů a podpisů, dále v lékařství (diagnostika) – vyhodnocování lékařských snímků (např. odhad velikosti tumoru prostaty), – detekce a ohodnocení v medicíně (např. detekce epileptických záchvatů), v ekonomice, marketingu a v expertních systémech v této oblasti – předpověď finančních časových řad (i chaotických) – predikce vývoje směnných kurzů, a v důležité oblasti jako je komprese dat, transformace a analýza signálu – komprese obrázků – diagnostika výsledků (např. rozpoznávání odběrových diagramů, ...). Tento výčet lze doplnit o nové mobilní aplikace pro překládání mluveného textu, rozpoznávání obličejů při identifikaci osob a využití přetrénování sítě (neuronová síť Caffe v projektu Deep Dream) k tvorbě art grafiky.

Dopředná (feed forward) neuronová síť



Obr. 1.10: Dopředná (feed forward) neuronová síť [113]

Nejrozšířenější způsob propojení neuronů se sigmoidní aktivační funkcí jsou vícevrstvé dopředné neuronové sítě. Tyto vícevrstvé sítě (MLP – Multilayer Perceptron) používají k učení Backpropagation (BPE – Back Propagation Errors)- algoritmus zpětného šíření chyby. Podmínkou je definice biasu ve výstupní a vnitřní vrstvě neuronů. Bias odpovídá váhové hodnotě přiřazené spojení mezi daným neuronem a fiktivním neuronem, jehož aktivace je vždy 1. Mezi dvěma sousedními vrstvami je úplné propojení neuronů, tedy každý neuron nižší vrstvy je spojen se všemi neurony vrstvy vyšší. Minimální počet vrstev je tři, vstupní, výstupní a alespoň jedna vnitřní vrstva.

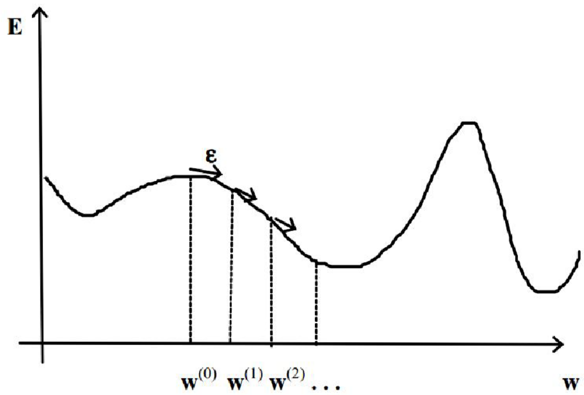
Metoda používaná pro adaptaci neuronové sítě nad danou trénovací množinou se nazývá backpropagation, nebo-li metoda zpětného šíření. Tato metoda adaptace spočívá v opačném šíření informace směrem od vrstev vyšších k vrstvám nižším. Během adaptace neuronové sítě metodou backpropagation jsou srovnávány vypočítané aktivace y_k s definovanými výstupními hodnotami t_k pro každý neuron ve výstupní vrstvě a pro každý tréninkový vzor. Na základě tohoto srovnání je definována chyba neuronové sítě, pro kterou je vypočítán faktor δ_k $(k = 1, \ldots, m)$, jež odpovídá části chyby, která se šíří zpětně z neuronu Y_k ke všem neuronům předcházející vrstvy majícím s tímto neuronem definované spojení. Podobně lze definovat i faktor δ_j ($j = 1, \ldots, p$), který je částí chyby šířené zpětně z neuronu vnitřní vrstvy Z_j ke všem neuronům vstupní vrstvy, jež mají s tímto neuronem definované spojení. Úprava váhových hodnot w_{jk} na spojeních mezi neurony vnitřní a výstupní vrstvy závisí na faktoru δ_k a aktivacích z_j neuronů Z_j ve vnitřní vrstvě. Úprava váhových hodnot v_{ij} na spojeních mezi neurony vstupní a vnitřní vrstvy závisí na faktoru δ_j a aktivacích x_i neuronů X_i ve vstupní vrstvě. Pro aktivační funkci pak požadujeme následující vlastnosti: musí být spojitá, diferencovatelná a monotónně neklesající. Nejčastěji používanou aktivační funkcí je proto standardní (logická) sigmoida a hyperbolický tangens. Chyba sítě $E(w)$ je vzhledem k tréninkové množině definována jako součet parciálních chyb sítě $E_l(w)$ vzhledem k jednotlivým tréninkovým vzorům a závisí na konfiguraci sítě w:

$$E(w) = \sum_{l=1}^{q} E_l(w),$$

kde parciální chyba $E_l(w)$ sítě pro l-tý tréninkový vzor ($l = 1, \ldots, q$) je úměrná součtu mocnin odchylek skutečných hodnot výstupu sítě pro vstup l-tého tréninkového vzoru od požadovaných hodnot výstupů u tohoto vzoru:

$$E_l(w) = \frac{1}{2} \sum_{k \in Y} (y_k - t_k)^2.$$

Cílem adaptace je dosažení globálního minima u chyby sítě, avšak tato chyba sítě přímo závisí na komplikované nelineární složené funkci vícevrstvé sítě, proto představuje tento cíl netriviální optimalizační problém. Pro jeho řešení se v základním modelu používá nejjednodušší varianta gradientní metody, která vyžaduje diferencovatelnost chybové funkce. Názorně ukazuje metodu obrázek:



Obr. 1.11: Gradientní metoda adaptace neuronové sítě [16]

Na obrázku je schematicky znázorněna chybová funkce $E(w)$ tak, že konfigurace, která představuje mnohorozměrný vektor vah w, se promítá na osu x. Chybová funkce určuje chybu sítě vzhledem k pevné tréninkové množině v závislosti na konfiguraci sítě. Při adaptaci sítě hledáme takovou konfiguraci, pro kterou je chybová funkce minimální. Začneme s náhodně zvolenou konfigurací $w^{(0)}$, kdy odpovídající chyba sítě od požadované funkce bude pravděpodobně velká. Při adaptaci sestrojíme v tomto bodě $w^{(0)}$ ke grafu chybové funkce tečný vektor (gradient) $\frac{\partial E}{\partial w}\left(w^{(0)}\right)$ a posuneme se ve směru tohoto vektoru o ε. Pro dostatečně malé ε tak získáme novou konfiguraci $w^{(1)} = w^{(0)} + \Delta w^{(1)}$, pro kterou je chybová funkce menší než pro původní konfiguraci $w^{(0)}$, tj. $E(w^{(0)}) \geq E(w^{(1)})$. Celý proces konstrukce tečného vektoru opakujeme pro $w^{(1)}$ a získáme tak $w^{(2)}$ takové, že $E(w^{(1)}) \geq E(w^{(2)})$, dokud se limitně nedostaneme do lokálního minima chybové funkce. Ve vícerozměrném váhovém prostoru tento postup přesahuje naši představivost. I když při vhodné volbě koeficientu učení (α) tato metoda vždy konverguje k nějakému lokálnímu minimu z libovolné počáteční konfigurace, není vůbec zaručeno, že se tak stane v reálném čase. Obvykle je tento proces časově velmi náročný i pro malé topologie vícevrstvé sítě (desítky neuronů). Pro dosažení globálního minima je pak třeba tuto funkci modifikovat (náhodné vnášení šumu do konfigurace sítě apod.) Ačkoliv vlastní popis učícího algoritmu backpropagation je formulován pro klasický von neumannovský model počítače, přesto je zřejmé, že jej lze implementovat distribuovaně. Pro každý tréninkový vzor probíhá nejprve aktivní režim pro jeho vstup tak, že informace se v neuronové síti šíří od vstupu k jejímu výstupu. Potom na základě externí informace učitele o požadovaném výstupu, tj. o chybě u jednotlivých vstupů, se počítají parciální derivace chybové funkce tak, že signál se šíří zpět od výstupu ke vstupu. Výpočet sítě při zpětném chodu probíhá sekvenčně po vrstvách, přitom v rámci jedné vrstvy může probíhat paralelně.

Problémem modelu vícevrstvé neuronové sítě s adaptačním algoritmem backpropagation je (kromě minimalizace chybové funkce) volba vhodné topologie pro řešení konkrétního praktického problému. Nejsou známy vztahy mezi vstupy a výstupy, které by se daly využít při návrhu speciální architektury. Používá se vícevrstvá topologie s jednou nebo dvěma vnitřními vrstvami za předpokladu, že učící algoritmus backpropagation zobecní příslušné vztahy z tréninkové množiny ve vahách jednotlivých spojů mezi neurony. Otázkou jsou počty neuronů ve vnitřních vrstvách. Počet neuronů vstupní vrstvy N_{in} je dán počtem vstupů a výstupní vrstvy počtem výstupů neuronové sítě. Počty neuronů N_k ve skrytých vrstvách nelze stanovit přesně, pro praktické aplikace však stačí většinou dodržet podmínku $N_k \geq \min(N_{out}, N_{in})$, kde N_{out} je počet neuronů výstupní vrstvy. Taktéž neexistuje ani exaktní pravidlo pro stanovení počtu úrovní neuronové sítě. Pro většinu praktických úloh jsou postačující tři úrovně: vstupní, jedna skrytá a výstupní, přičemž vstupní úroveň pouze opakuje vstupní signály ($y_i^1 = x_i$), v případně potřeby ji lze využít pro transformaci vstupních hodnot. Problematikou jejich aproximativních vlastností se zabývalo mnoho autorů. Nejdůležitějším jsou zřejmě důsledky Kolmogorovovy věty o reprezentaci spojitých funkcí více proměnných pomocí spojitých funkcí jedné proměnné. Nejprve je třeba definovat některé vybrané pojmy:

- Prostor všech spojitých funkcí nad kompaktním prostorem[2] X (označovaný $C(X)$) se stejnoměrnou normou:

$$||f|| = \sup_{x \in X} |f(x)|.$$

- Prostor $\mathcal{L}_p$ funkcí, pro něž je $|f|^p$ integrovatelná na množině X, s tzv. $\mathcal{L}_p$ normou:

$$||f|| = \left[\int_{x \in X} |f(x)|^p dx \right]^{\frac{1}{p}}.$$

V prostoru $\mathbb{R}^n$ často za kompaktní $X \subset \mathbb{R}^n$ uvažujeme jednotkovou krychli $\mathcal{I}^n$.

Věta 1. *Buďte $n \geq 2$ přirozené číslo a $f(x_1, \ldots, x_n)$ spojitá funkce $f : \mathcal{I}^n \to \mathbb{R}$. Pak existují spojité funkce jedné proměnné ψ_q a ϕ_{pq} $(p = 1, \ldots, n; q = 1, \ldots, 2n + 1)$ tak, že platí:*

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \psi_q \left(\sum_{p=1}^{n} \phi_{pq}(x_p) \right).$$

V alternativním zápisu potom takto:

Každou vícerozměrnou reálnou spojitou funkci $f(x_1, \ldots, x_n)$ n-proměnných lze přesně vyjádřit jako lineární kombinaci konečného počtu spojitých nelineárních funkcí jedné proměnné:

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \psi_q \left(\sum_{p=1}^{n} \phi_{pq}(x_p) \right)$$

kde ψ, ϕ jsou nelineární funkce jedné proměnné.

Zde jednotková krychle $\mathcal{I}^n$ představuje kompaktní podprostor $\mathbb{R}^n$. Prostor všech spojitých funkcí nad kompaktním prostorem $X \subset \mathbb{R}^n$ pak označujeme jako $C(X)$, s vyznačením dimenze ve vlastních výsledcích potom jako funkcionální krychle.

Obecnější modifikace této věty byla představena Spredlerem a Lorentzem [80, 113] V této variantě se zápis změnil na:

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{2n+1} \psi \left(\sum_{p=1}^{n} \lambda^{pq} \phi_q(x_p) \right). \tag{1.1}$$

[2]Kompaktní množina, nebo také kompaktní prostor, je taková množina bodů topologického prostoru, že z každého jejího pokrytí otevřenými množinami lze vybrat pokrytí konečné. Tato definice v topologii zobecňuje a formalizuje intuitivní představu konečného objemu. V Eukleidovských prostorech jsou kompaktní množiny právě omezené a uzavřené podmnožiny.[87]

Zde jedna funkce ψ nahradila všechny ψ_q a funkce ϕ_{pq} jsou nahrazeny výrazem $\lambda^{pq}\phi_q$, kde λ je reálná konstanta a ϕ_q jsou monotónně rostoucí funkce. Navíc, ϕ_q nezávisí na funkci f, kterou vyjadřujeme, ale jsou pro danou dimenzi stejné.

Kolmogorovovu větu lze využít i pro běžné perceptronové sítě, pokud nebudeme trvat na přesném vyjádření funkce, ale spokojíme se s její libovolně přesnou aproximací. V tom případě dostaneme síť se dvěma skrytými vrstvami perceptronů, kde jednotky v první, resp. druhé skryté vrstvě budou aproximovat funkce ψ, resp. ϕ.

Pro další popis funkcionality neuronové sítě v souvislosti s aproximací v Kolmogorově větě je užitečné uvést definici topologického prostoru:

Definice 1. *[87] Topologickým prostorem nazveme množinu X společně se systémem τ podmnožin X, splňující následující axiomy:*

- $\emptyset \in \tau$, $X \in \tau$,
- *sjednocení libovolného počtu (tj. konečného, spočetného i nespočetného) množin z τ leží v τ,*
- *průnik konečného počtu množin z τ leží v τ.*

Systému τ říkáme topologie[3] na X. Množiny v τ pak nazveme otevřené množiny, jejich doplňky v X uzavřené množiny.

Konkrétní topologický prostor bývá často označován jako (X, τ).

Poznámka 1. Kolmogorova věta vznikla jako součást řešení třinácté části Hilbertovy přednášky [39] Matematické problémy, přednesené na 2. Mezinárodním kongresu matematiků r. 1900 v Paříži, která má název „Nemožnost řešení obecné rovnice sedmého stupně pomocí funkcí dvou proměnných“. V této části D. Hilbert předkládá problém inspirovaný nomogramem[4], který se týká jedné z nejstarších úloh matematiky – řešitelnosti algebraických rovnic. Obecná teorie aproximace je definována pomocí základních pojmů matematické topologie:

- Podmnožina $V \subset X$ topologického prostoru (X, τ) je otevřená, když ke každému $x \in V$ existuje okolí obsažené ve V. Množina $F \subset X$ je uzavřená, je-li $X \setminus F$ otevřená.
- Pro $A \subset X$ definujeme uzávěr $\overline{A}$ jako nejmenší uzavřenou množinu obsahující A. Dále říkáme, že $A \subset X$ je hustá v X, je-li $\overline{A} = X$, tedy uzávěr této množiny je již celý prostor.

Topologický prostor X je Hausdorffův, když pro každou dvojici různých bodů $x_1, x_2 \in X$ existují otevřené množiny $U_1, U_2 \subset X$ takové, že $x_1 \in U_1, x_2 \in U_2$ a $U_1 \cap U_2 = \emptyset$. Platí, že každý prostor s topologií indukovanou metrikou je Hausdorffův.

[3]Pojem topologie se také používá pro matematickou disciplínu.

[4]Nomogram je speciální graf, který umožňuje provádění výpočtů pomocí jednoduchých geometrických konstrukcí a čtením přímo v tomto grafu. Nejjednodušším typem nomogramu je graf funkce v kartézské soustavě souřadnic [1].

Definice 2. *Nechť U je třída funkcí, T její podmnožina a ρ metrika na U. Třídu T nazveme aproximátorem vzhledem k (U, ρ), je-li T hustá v U (vzhledem k topologii indukované metrikou ρ). Je-li třída funkcí T aproximátorem vzhledem ke třídě spojitých reálných funkcí (nebo $\mathcal{L}_p$ funkcí), nazveme ji univerzálním aproximátorem.*

Využití Kolmogorovovy věty v neuronových sítích navrhl v roce 1987 Hecht-Nielsen [37], který si všiml, že vzorec (1.1) připomíná zápis funkce realizované sítí se třemi vrstvami. Vstupní vrstvu takové sítě tvoří n vstupních jednotek, ve skryté vrstvě je $2n + 1$ jednotek a jedna jednotka je výstupní. Mezi vstupní a skrytou vrstvou jsou přechodové funkce počítající hodnoty $\lambda^{pq}\phi_q$, přechodovou funkcí mezi skrytou a výstupní vrstvou je ψ. Tento poznatek byl později prokázán právě pro aproximační funkci neuronové sítě.

Nejvíce výsledků o univerzální aproximaci neuronových sítí se týká perceptronových sítí s jednou skrytou vrstvou. Tyto sítě jsou v praxi často používaným modelem a jedna skrytá vrstva je v jistém smyslu minimální architektura zachovávající rysy vícevrstvých perceptronů. Motivací pro výzkum aproximačních vlastností možnosti modelu aproximovat nejrůznější funkce. Následující poznámka ukazuje jejich aproximační možnosti při jednoduché základní struktuře:

Poznámka 2.

- Je-li aktivační funkce omezená a nekonstantní, pak perceptronová síť s jednou skrytou vrstvou muže aproximovat každou funkci z Λ_p s libovolnou přesností za předpokladu, že je k dispozici dostatečně mnoho jednotek ve skrytě vrstvě.
- Je-li aktivační funkce spojitá, omezená a nekonstantní na kompaktní množině $X \subset R^n$, pak může dopředná síť s jednou skrytou vrstvou ϕ aproximovat každou funkci $g \in C(X)$ libovolně přesně (ve smyslu $|g - \phi| < \varepsilon$ pro libovolně malé ε), je-li k dispozici dostatečně mnoho neuronů ve skryté vrstvě.

Tento problém organizační dynamiky a optimálního počtu vrstev a neuronů úzce souvisí s adaptací a **generalizací** neuronové sítě, schopností vystihnout správný výsledek na nových datech. S tím úzce souvisí princip rozdělení datového souboru, který neuronové sítě implicitně používají:

- Trénovací množina – náhodně vybraná část dat, která slouží pro učení sítě.
- Testovací množina – další část dat sloužící k zastavení trénovaní, aby nedošlo k přeučení sítě.
- Validační množina – zbytek dat, na kterém ověříme konečnou kvalitu modelu. Jde o data, která dosud model k dispozici neměl.

Při učení pomocí algoritmu backpropagation se příliš malá síť zastaví v nějakém mělkém chybovém lokálním minimu, kdy je třeba doplnit další neurony. Příliš bohatá architektura sice při učení mnohdy umožní nalézt globální minimum chybové funkce, ale příliš zobecňuje tréninkové vzory včetně jejich nepřesností a chyb a pro nenaučené vzory dává chybné výsledky, tj. špatně generalizuje. Tomuto přesnému zapamatování tréninkové množiny bez zobecnění zákonitostí v ní obsažených se říká **přeučení (overfitting)**. Příčinou může být malý rozsah trénovací množiny nebo přílišná komplexnost systému (např. příliš mnoho skrytých neuronů v neuronové síti). Řešením je zvětšení trénovací množiny, snížení složitosti systému nebo různé techniky regularizace jako je zavedení náhodného šumu (což v zásadě odpovídá rozšíření trénovací množiny), zavedení omezení na parametry systému, které v důsledku snižuje složitost popisu naučené funkce, nebo předčasné ukončení (průběžné testování na validační množině a konec učení ve chvíli, kdy se chyba na této množině dostane do svého minima). Další možnou chybou je **přetrénování sítě (overlearning)**. V ideálním případě konverguje celková chyba na trénovací množině k nule (= stav stoprocentní natrénovanosti sítě na trénovací množinu), zatímco chyba sítě na testovací množině nejprve klesá a následně roste. Stav minima chyby na testovací množině je optimálním okamžikem pro ukončení učení, neboť od tohoto okamžiku dál síť ztrácí svou schopnost zobecňovat a je příliš fixována na trénovací množinu – stává se přetrénovanou.

1.3 Prostředky algebraizace vícevrstvé dopředné neuronové sítě

Matematické modelování představuje transformaci z aplikační oblasti do oblasti matematicky řešitelné formulace umožňující teoretickou a numerickou analýzu problému. Po uchopení podstaty problému spočívá další postup v analýze získaných poznatků, objevení dalších souvislostí a nalezení jejich možného uplatnění v praxi. K tomuto účelu se používají modelovací časové funkce. Dopředná neuronová síť, která byla cílem výzkumu této práce, je často předmětem výzkumu modelování z důvodu jejího častého použití a vývoje aplikací. Na neuronovou síť je možné aplikovat poznatky z analýzy a filtrace signálů, z oblasti zobrazení vektorových prostorů, z pohledu topologie z oblasti grafů a také poznatky z teorie systémů. Tyto poznatky byly zkoumané v řadě obsáhlých publikací [16, 20, 34, 113, 114]. V řadě z nich se objevují aproximace neuronové sítě, například v knize od C. M. Bishopa [15] (Neural Networks for Pattern Recognition) je uveden modelový příklad aproximace pomocí polynomů včetně možných problémů při volbě aproximační funkce. Tato práce, jak již bylo řečeno v úvodu, je motivovaná ideou zobecnění funkce časově proměnného neuronu a pomocí analogie s lineárními diferenciálními operátory a jejich algebraickými strukturami tvořit a zkoumat algebraické struktury těchto neuronů na podkladě funkce dopředné neuronové sítě – vícevrstvého perceptronu MLP. Pro tento účel bylo využito poznatků z oblasti lineárních diferenciálních operátorů, algebraických struktur a hyperstruktur a orientovaných grafů.

Vybrané poznatky o algebraických strukturách a hyperstrukturách

Při modelování a zkoumání funkce dopředné neuronové sítě byly uplatněny poznatky získané zkoumáním vlastností hyperstruktur. Multistruktury nazývané také hyperstruktury patří k části moderní algebry, která je stále cílem výzkumu a aplikace. Pojem hypergrupy byl zaveden již v roce 1934, na 8. kongresu skandinávských matematiků, kde Fréderic Marty [82] poprvé představil tuto teorii, a na něho navázali Dresher a Ore [27]. Teorie hyperstruktur a zejména teorie hypergrup je obsažena v několika oblastech matematiky jako je geometrie (deskriptivní, sférická, projektivní) [26], grafy a hypergrafy a binární relace, uspořádané množiny a zejména svazy, uspořádané algebraické struktury, automaty, kryptografie, kódy, obecné systémy, umělá inteligence, polygroupy (aplikované v kombinatorice), pravděpodobnost, fuzzy množiny a některé další aplikace a speciální konstrukce. Teorii hyperstruktur je věnováno množství publikací, včetně [6, 41, 42, 43], kde lze najít další zajímavé aplikace. Jedním z cílů výzkumu neuronové sítě bylo nalézt aplikaci těchto poznatků při modelování funkce neuronové sítě. Pro aplikaci konceptu teorie hyperstruktur je užitečné vycházet ze základních poznatků [53, 62]:

Definice 3. *Nechť H je neprázdná množina a $*$ je zobrazení kartézkého součinu $H \times H$ do systému všech neprázdných podmnožin množiny H (běžně označované $\mathcal{P}^*(H)$). Dvojice $(H, *)$ se nazývá hypergrupoid. Pro každé dvě neprázdné podmnožiny A a B množiny H a $x \in H$ definujeme:*

$$A * B = \bigcup_{a \in A, b \in B} a * b, \quad A * x = A * \{x\}, \quad x * B = \{x\} * B.$$

Definice 4. *Hypergrupoid $(H, *)$ se nazýva polohypergrupa, když pro všechna $a, b, c \in H$ platí $(a * b) * c = a * (b * c)$, tím rozumíme, že*

$$\bigcup_{u \in a*b} u * c = \bigcup_{v \in b*c} a * v$$

Hypergrupoid $(H, *)$ se nazývá kvazihypergrupa, když pro každé $a \in H$ platí $a * H = H = H * a$. Tato podmínka se nazývá reprodukční zákon (axiom).

Pojem hypergrupy popisuje následující definice [65]:

Definice 5. *Hypergrupa je asociativní hypergrupoid $(H, *)$, tj. $a. * (b * c) = (a * b) * c$ pro libovolnou trojici $a, b, c \in H$, splňující reprodukční axiom: $a * H = H * a = H$ pro libovolný prvek $a \in H$.*

Podhypergrupoid hypergrupy $(H,*)$ je dvojice $(S,*)$, kde $S*S \subset S \subset H$. Poznamenejme, že relace incidence neprázdných množin A, B, tedy $A \cap B \neq \emptyset$, se v literatuře o hyperstrukturách obvykle označuje $A \approx B$.

Při konstrukci algebraických struktur na bázi neuronů dopředné sítě je vhodné připomenout pojem spojnicový prostor. Spojnicové prostory byly zavedeny Walterem Prenowitzem a Jamesem Jantosciakem [102] při zkoumání jistých geometricky motivovaných nekomutativních struktur.

Definice 6. *Hypergrupa $(H,*)$ se nazývá transpoziční hypergrupa nebo také (nekomutativní) spojnicový prostor, jestliže splňuje axiom transpozice: Pro každou čtveřici $a,b,c,d \in H$ ze vztahu $b \setminus a \approx c/d$ plyne $a*d \approx b*c$, kde množiny $b \setminus a = \{x \in H; a \in b*x\}$, $c/d = \{x \in H; c \in x*d\}$ se nazývají levá a pravá extenze (někdy též levý a pravý zlomek).*

Důležitá je v problematice neuronové sítě otázka uspořádání a tedy i uspořádaných algebraických struktur. Tyto otázky společně s některými poznatky teorie grafů vstupují do širší problematiky optimalizace funkčnosti neuronové sítě.

Definice 7. *Uspořádanou grupou rozumíme trojici $(G,\bullet,\leq)$, kde $(G,\bullet)$ je grupa a binární relace $\leq$ je uspořádání na G takové, že pro libovolnou trojici $x,y,z \in G$ plyne z vlastnosti $x \leq y$ také $x \bullet z \leq y \bullet z$, $z \bullet x \leq z \bullet y$.*

V uspořádané grupě je symbolem $[a)_{\leq}$ označen hlavní konec generovaný prvkem $a \in G$, definovaný takto: $[a)_{\leq} = \{x \in G; a \leq x\}$. Tvrzení o vlastnostech hypegrupy, definované pomocí výše uvedeného pojmu, označujeme jako „koncové lemma". Následující věta, které je rozhodující pro další konstrukci struktur umělých neuronů, je dokázáno v [55, 57] (česká verze je dokázána v [51], s. 146, 147):

Věta 2. *Nechť máme uspořádanou grupu $(G,\bullet,\leq)$. Pro každou dvojici prvků $a,b \in G$ definujeme hyperoperaci $*$ na množině G takto: $a*b = [a \bullet b)_{\leq}$. Potom $(G,*)$ je hypergrupa, která je komutativní, právě když grupa $(G,\bullet)$ je komutativní.*

Nyní nechť A^* je volný monoid slov nad nad libovolnou (neprázdnou) abecedou A; předpokládejme, že e znamená prázdné slovo.

V souladu s [28], [59] a dalšími publikacemi, automatem máme na mysli trojici $\mathbb{A} = (S,A,\delta)$, kde S,A jsou libovolné množiny $(A \neq 0)$, které se – v daném pořadí – nazývají množinou stavů (nebo stavovou množinou), množinou vstupních symbolů (nebo vstupní abecedou) a $\delta : S \times A \to S$ je zobrazení, nazývané přechodová funkce, které splňuje tyto dvě podmínky:

- $\delta(s,e) = s$ pro libovolný stav $s \in S$, (axiom identity).
- $\delta(s,ab) = \delta(\delta(s,a),b)$ pro libovolný stav $s \in S$ pro libovolný pár slov $a,b \in A^*$ (axiom homomorfismu). Je třeba poznamenat, že přechodová funkce $\delta : S \times A^* \to S$ je obvyklá extenze funkce příštího stavu $\delta_{\mathbb{A}}^{(0)} : S \times A \to S$.

Automat $\mathbb{A}_1 = (S_1, A, \delta_1)$ je považován za podautomat automatu $\mathbb{A} = (S, A, \delta)$ jestliže $S_1 \subseteq S, \delta(s, a) \in S_1$ pro libovolný stav $s \in S_1$ a libovolné slovo $a \in A^*$ a dále $\delta_1 = \delta \uparrow (S \times A^*)$, tj. δ_1 je restrikce přechodové funkce δ na $S_1 \times A^*$.

Pro $\emptyset \neq T \subseteq S$ označujeme $\delta(T, A^*) = \{\delta(t, a); t \in T, a \in A^*\}$. (v [7], [22] je tento operátor označen pouze jako δ, tj. $\delta(T)$ znamená výše uvedenou množinu v citovaných zdrojích). Jestliže T je singleton, tj. $T = \{t\}$, píšeme $\delta(t, A^*)$ místo $\delta(\{t\}, A^*)$.

Nechť $\mathbb{A} = (S, A, \delta)$ je automat. Definujeme binární operaci na množině stavů S množiny A takto:

$$s \circ t = \delta(s, A^*) \cup \delta(t, A^*).$$

pro libovolný pár stavů $s, t \in S$. Poněvadž A^* obsahuje prázdné slovo e, získáme $s, t \in s \circ t$, tudíž $s \circ t. \neq \emptyset$ a také komutativita této hyperoperace je rovněž zřejmá. Dále, $s^2 = \delta(s, A^*)$ takže $s \circ t = s^2 \cup t^2$ a platí:

$$\begin{gathered}
s^3 = s \circ s^2 = s \circ \delta(s, A^*) = \textstyle\bigcup_{t \in \delta(s, A^*)} s \circ t = \delta(s, A^*) \cup \bigcup_{t \in \delta(s, A^*)} \delta(t, A^*) = \\
= s^2 \cup \delta(\delta(s, A^*),\, A^*) = s^2 \cup \{\delta(\delta(s, a), b); a \in A^*),\, b \in A^*)\} = \\
= s^2 \cup \{\delta(s, ab); ab \in A^*A^* = A^*\} = s^2 \cup \{\delta(s, c); c \in A^*\} = s^2 \cup \delta(s, A^*) = \\
= s^2 \cup s^2 = s^2
\end{gathered}$$

tj. $s \in s^2 = s^3$ pro libovolný stav $s \in S$. V důsledku toho je hyperoperace $\circ$ asociativní a získáváme poznatek, že $(S, \circ)$ je kvazi-uspořádaná hypergrupa. Totéž také vyplývá z faktu, že $s \circ t = \{u \in S; (s, u) \in r$ nebo $(t, u) \in r\} = r(s) \cup r(t)$, kde r je tranzitivní uzávěr relace ν na množině stavů $\mathbb{A}((s, t) \in \nu$ jestliže $\delta(s, a) = t$ pro nějaký vstupní symbol $a \in A)$, který byl užit a studován M. W. Warnerem v [118]. Právě definovaná kvazi-uspořádaná hypergrupa $(s, \circ)$ je nazývaná stavová hypergrupa automatu $\mathbb{A} = (S, A, \delta)$ označená také jako $\mathbb{H}(\mathbb{A})$. Pro další výzkum jsou důležité také typy kompatibility binárních relací.

Typy kompatibility binárních relací

Pro další výzkum algebraických struktur je důležitý koncept relační kompatibility (slučitelnosti) na hypergrupách $(H, \cdot)$: **Substituční vlastnost - SP**

Definice 8. ***Substituční vlastnost - SP [65, 44]*** *Řekneme, že binární relace $R \subseteq H \times H$ má substituční vlastnost*

- ***prvního typu (SP-1)*** *jestliže pro libovolnou čtveřici $a, b, c, d \in H$, pro niž platí aRb, cRd máme $(a \cdot c)\ \overline{R}\ (b \cdot d)$*[5]*, tj. pro každé $x \in a \cdot c$ zde je nějaké*

[5]Relace $\overline{R}$ je definována takto: $X\ \overline{R}\ Y$ pro $X, Y \subseteq H$, jestliže pro $\forall x \in X$, $\exists y \in Y$ s vlastností xRy a pro $\forall y' \in Y\ \exists x' \in X$ s vlastností $x'Ry'$. Dále $X\ \overline{\overline{R}}\ Y$ znamená, že pro $\forall x \in X$ a $\forall y \in Y$ platí xRy.

$y \in b \cdot d$ takže platí xRy a naopak, tak že pro libovolné $y \in b \cdot d$ zde existuje nějaké $x \in a \cdot c$ s vlastností xRy,

- ***druhého typu(SP-2)*** *jestliže pro libovolnou čtveřici $a, b, c, d \in H$ a zároveň aRb, cRd platí $R(a \cdot c) = R(b \cdot d)$ (zde, jak je uvedeno výše, $R(X) = \{y \in H | \exists x \in X$ s vlastností $xRy\}$),*
- ***třetího typu(SP-3)*** *jestliže $a, b, c, d \in H$, aRb, cRd implikuje $(a \cdot c) \overline{\overline{R}} (b \cdot d)$, tj. pro libovolný pár $(x, y) \in (a \cdot c) \times (b \cdot d)$ pak máme xRy.*

Parametrizovatelnost pomocí endomorfismů. O relaci $R \subseteq H \times H$ řekneme, že je parametrizovatelná pomocí množiny Φ zobrazení definičního oboru $\mathcal{D}R$ do H jestliže pro libovolný pár $(a, b) \in R$ zde existuje $f \in \Phi$ tak, že platí $f(a) = b$ a $\{(x, f(x)) \mid x \in \mathcal{D}R\} \subseteq R$, tj. $R = \cup\{grf | f \in \Phi\}$, kde binární relace $grf = \{(a, f(a) | a \in \mathcal{D}R\})$ je grafem zobrazení $f : \mathcal{D}R \to H$. Kvůli stručnosti píšeme pouze f místo grf. O relaci $R \subseteq H \times H$ řekneme, že je *parametrizovatelná pomocí množiny endomorfismů (dobrých endomorfismů)* nebo krátce *parametrizovatelná pomocí endomorfismů (dobrých endomorfismů)* jestliže je parametrizovatelná pomocí $\Phi \subseteq End(H, \cdot)$ ($\Phi \subseteq GEnd(H, \cdot)$).

Zachování množinového systému. Nechť $\emptyset \neq \mathfrak{G}_H \subseteq \mathcal{P}(H)$. Relace $R \subseteq H \times H$ je nazývána $\mathfrak{G}_H$-zachovávající ($\mathfrak{G}_H$-reflektující nebo $\mathfrak{G}_H$-vracející), jestliže zachovává (vrací) členy $\mathfrak{G}_H$, tj. $X \in \mathfrak{G}_H$ implikuje $R(X) \in \mathfrak{G}_H$ ($R^-(X) \in \mathfrak{G}_H$, tj. R^- je $\mathfrak{G}_H$-zachovávající). Zde R^- je inverzní relace k R (cf. [74], která je také označena jako R^{-1}, $R^{\cup}$ jinde a nazývána jako inverze nebo konverze relace k R, případně opačná relace k R.).

Věta 3. *Nechť $(G, \cdot)$,$(H, \cdot)$ jsou kvazi-uspořádané hypergrupy, $R \subseteq G \times H$ je celodoménová relace parametrizovatelná homomorfismy hypegrupy $(G, \cdot)$ do hypergrupy $(H, \cdot)$. Pak platí následující tvrzení:*

1. *Relace R je $\mathfrak{C}Q$-reflektující.*
2. *Pro libovolný pár $x, y \in G$ máme $R(x \cdot y) \subseteq R(x) \cdot R(y)$.*
3. *Pro libovolnou neprázdnou podmnožinu $X, Y \subseteq G$ platí $R(X \cdot Y) \subseteq R(X) \cdot R(Y)$ (zejména $R(X^2) \subseteq [R(X)]^2$).*

Věta 4. *Nechť $(G, \cdot)$,$(H, \cdot)$ jsou kvazi-uspořádané hypergrupy, $R \subseteq G \times H$ je celodoménová relace parametrizovatelná dobrými homomorfismy hypegrupy $(G, \cdot)$ do hypergrupy $(H, \cdot)$. Pak R splňuje (1), (2), (3) z Věty 3 a navíc následující podmínky:*

1. *Relace R zachovává multiplikativně uzavřené podmnožiny, tj. je Q-zachovávající.*
2. *Pro libovolný pár $(a, b), (c, d) \in R \times R$ a libovolný pár $(x, y) \in (a \cdot c) \times (b \cdot d)$ platí $R(x) \cap (b \cdot d) \neq \emptyset \neq (a \cdot c) \cap R^-(y)$.*

Důkazy uvedených vět a další zdroje lze nalézt v článku [44].

W-kompatibilita W-kompatibilita (kompatibilita Wallaceho typu) je kompatibilita binárních relací mezi algebraickými strukturami, studovaná A. D. Wallacem v roce 1960. Je motivovaná kompatibilitou na topologických prostorech.

Definice 9. *[63]* ***W–kompatibilita.*** *Relace $R \subseteq H \times H$ se nazývá kompatibilní ve Wallaceově smyslu nebo zkráceně W–kompatibilní s hypergrupou $(H, \cdot)$, jestliže pro každou podmnožinu $X \subseteq H$ platí: $R^{-}(E_{*}(X)) \subseteq E_{*}(R^{-}(X))$. Symbolem R^{-} je označena relace inverzní k relaci R. Přitom $E_{*}(\emptyset) = \emptyset, E_{*}(X) = X \cup X^2 \cup \ldots \cup X^n, n \in N$.*

Algebraizace pomocí lineárních diferenciálních operátorů

Při bližším pohledu na funkcionalitu dopředné neuronové sítě je ze vztahu (1.1) zřejmá role přenosové funkce. Při její vhodné volbě při modelování algebraických struktur je patrná podobnost s lineárním diferenciálním operátorem.

Jestliže ke každé modelovací funkci nalezneme její první a druhou derivaci a poté z první a druhé derivace sestavíme lineární diferenciální rovnici druhého řádu, můžeme z takto sestavené diferenciální rovnice získat lineární diferenciální operátor druhého řádu. Obecně lze tento přístup rozšířit do dimenze n pro lineární diferenciální operátory n-tého řádu. Tyto operátory uvažujeme jako prvky množin, na kterých lze zavést rozličné binární hyperoperace a studovat jejich vlastnosti. Tento postup užití lineárních diferenciálních operátorů získaných z lineárních diferenciálních rovnic vychází z práce profesora Otakara Borůvky a jeho následovníků, jako je např. profesor Neuman [58, 59, 60, 61], ale i další autoři studují z hlediska algebraického přístupu lineární diferenciální operátory. V těchto výsledcích lze najít motivaci nejen při hledání řešení diferenciálních rovnic.

Předpokládejme, že existuje zobrazení A z prostoru funkcí $\mathcal{F}_1$ do prostoru funkcí $\mathcal{F}_2$, a že funkce $f \in \mathcal{F}_2$, taková, že f je obrazem $u \in \mathcal{F}_1$, tj. $f = A(u)$. Diferenciální operátor je reprezentován jako lineární kombinace konečně generovaná u a jeho derivacemi vyššího stupně. Vezměme v úvahu poznatky z oblasti lineárních diferenciálních rovnic. Zde je použit tvar operátoru obyčejných lineárních diferenciálních rovnic v této formě:

$$L_n = \sum_{k=0}^{n} p_k(x) D^k, \tag{1.2}$$

kde podle prací britského matematika Olivera Heaviside $D^k = \frac{d^k}{dx^k}$, $p_k(x)$ je spojitá funkce na nějakém intervalu $J \subset \mathbb{R}$, $k = 0, 1, \ldots, n-1$, $p_n(x) \equiv 1$, to znamená, že $L_n(y) = 0$ je lineární homogenní diferenciální rovnice ve tvaru:

$$y^{(n)}(x) + \sum_{k=0}^{n-1} p_k(x) y^{(k)}(x) = 0.$$

Tyto operátory tedy tvoří levé strany obyčejných lineárních diferenciálních rovnic n-tého řádu.

2 Cíle disertační práce

Předložená disertační práce se zabývá parametrizací přechodových relací strukturovaných objektů v oblasti hypergrup a automatů s aplikací v oblasti neuronových sítí s vhodnou strukturou, aktivační funkcí, s cíleným zjednodušením modelu. Ve výzkumném projektu, který popisuje disertační práce, je zkoumána problematika dopředné neuronové sítě s akcentem na algebraické struktury konstruované na bázi umělých neuronů. Oblast neuronových sítí je velice široká s mnoha aplikacemi a možnostmi jejich uplatnění. Proto je výzkumným polem disertační práce jen úzká vybraná součást problematiky neuronových sítí – algebraické struktury v jednoduché dopředné síti s aplikací časově proměnných neuronů. V souvislosti se strukturami a multistrukturami jsou studovány jejich vztahy v souvislosti s parametrizací relací. Je zkoumán pojem řešitelnosti, modely iterovaných umělých neuronů, akce grup a hypergrup časově proměnných neuronů, a také uspořádání zachovávající zobrazení v řadě hypergrup časově proměnných neuronů určitých vlastností. Získané poznatky mohou být použity při dalším rozvoji neuronových sítí. Hlavní cíl práce, kterým je prozkoumání vlastností algebraických struktur v oblasti neuronových sítí a jejich dalších vlastností, bude konkretizován pomocí následujících dílčích cílů:

1. Zkoumat možnosti parametrizace relací.
2. Navržení konstrukce algebraických struktur na vícevrstvé neuronové síti s vybranou přechodovou funkcí. Použití analogie s již dříve zkoumanými modely pro vytvoření grupy neuronů s využitím časově proměnných neuronů.
3. Zkoumat vlastnosti vytvořené neuronové algebraické struktury a doplnit výsledky týkající se řešitelnosti a dalších vlastností relací mezi strukturami a hyperstrukturami časově proměnných neuronů dané grupy.
4. Zkoumat akce struktur a hyperstruktur časově proměnných neuronů a modely iterovaných časově proměnných neuronů.
5. Zkoumat systém součástí neuronové sítě z pohledu teorie grafů a uspořádaných množin.

3 Vlastní výsledky

3.1 Parametrizace relací

3.1.1 Parametrizace relací pomocí morfismů

Tvrzení 1. *Nechť* $(G,*)$, $(H,\cdot)$ *jsou komutativní hypergrupy,* $R \subseteq G \times H$ *je celo-doménová relace parametrizovatelná pomocí homomorfismů hypergrupy* $(G,*)$ *do hypergrupy* $(H,\cdot)$. *Pak* $R(\langle X\rangle_H) \subseteq \langle R(X)\rangle$ *pro libovolnou neprázdnou podmnožinu* $X \subseteq G$ [1].

Důkaz. Předpokládejme $R = \bigcup\limits_{\gamma\in\Gamma} f_\gamma$ pro vhodný systém homomorfismů $f_\gamma : (G,*) \to (H,\cdot)$, $\emptyset \neq X \subseteq G$. Poněvadž $f_\gamma \subseteq R$ pro každé $\gamma \in \Gamma$ máme $f_\gamma(X) \subseteq R(X)$, což implikuje $\langle f_\gamma(X)\rangle \subseteq \langle R(X)\rangle$ na základě monotonie operátoru $\langle\ \rangle$ a tak $\bigcup\limits_{\gamma\in\Gamma}\langle f_\gamma(X)\rangle \subseteq \langle R(X)\rangle$. Pak platí

$$R(\langle X\rangle_H) = \left(\bigcup_{\gamma\in\Gamma} f_\gamma\right)(\langle X\rangle) = \bigcup_{\gamma\in\Gamma} f_\gamma(\langle X\rangle) \subseteq \bigcup_{\gamma\in\Gamma}\langle f_\gamma(X)\rangle \subseteq \langle R(X)\rangle.$$

□

Nyní vezměme v úvahu mono-unární algebru (H,f), tj. neprázdnou množinu H a zobrazení $f : H \to H$. Označme $\mathbb{N}_0 = N \cup \{0\} = \{0,1,2,...\}$ a definujme binární relaci na H takto:

$$R_f = \{(x, f^m(x)) | x \in H, m \in \mathbb{N}_0 - \{1\}\}.$$

Snadno se vidí, že R_f je kvazi-uspořádání na H což je uspořádání právě tehdy, když transformace množiny $f : H \to H$ má maximálně jednoprvkové cykly, tj. smyčky. KW-ekvivalence $\sim_f$ (Kuratowského–Whyburnova typu) na mono-unární algebře (H,f) je definovaná takto: $x \sim_f y \equiv f^n(y)$ pro nějaký pár m,n nezáporných celých čísel. Bloky této relace ekvivalence se nazývají orbity množinové transformace $f : H \to H$, podalgebry $(K, f \restriction K)$, kde $K \in H/\sim_f$ se nazývají komponenty mono-unární algebry (H,f).

Příklad. Tento jednoduchý příklad ukazuje, že parametrizovatelnost dobrými endomorfismy binární relace na hypergrupě neimplikuje substituční vlastnost. Vezměme v úvahu Peanovu algebru všech kladných celých čísel $(\mathbb{N}, f)$, kde $f(n) = n+1$ pro každé $n \in \mathbb{N}$. Označme $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ a definujme zobrazení $\varphi, \psi : \mathbb{N} \to \mathbb{N}$ jako $\varphi(1) = 1$, $\varphi(n) = n-1$ pro $n \geq 2$, $\psi(1) = \psi(2) = 1$ a $\psi(n) = n-2$ pro $n \geq 3$. Nyní definujme $C_f X = \bigcup\limits_{k\in\mathbb{N}_0} f^k(X)$, kde f^k je k-tá iterace f. Nyní pro libovolné

[1] $\langle X\rangle_H$ je uzávěrový operátor na H, který každé podmnožině X přiřazuje nosič podhypergrupy hypergrupy H, a to nejmenší vzhledem k množinové inkluzi,který obsahuje podmnožinu X.

$x, y \in \mathbb{N}$ definujeme $x \cdot y = C_f\{x, y\}$. Pak $(\mathbb{N}, \cdot)$ je komutativní částečně uspořádaná hypergrupa a relace $R = \varphi \cup \psi$ je parametrizovatelná pomocí endomorfismů $\varphi, \psi : (\mathbb{N}, \cdot) \to (\mathbb{N}, \cdot)$. Dále $(1, 1) \in R$, avšak $(f(1), f(1)) = (2, 2) \notin R$ takže R nemá SP vlastnost na mono-unární algebře $(\mathbb{N}, f)$. Podobným způsobem můžeme definovat relaci na $\mathbb{N}$, parametrizovatelnou pomocí endomorfismů hypergrupy $(\mathbb{N}, \cdot)$, která také nemá SP-1 na $(\mathbb{N}, \cdot)$.

Tvrzení 2. *Nechť H je neprázdná množina a $f : H \to H$ je surjektivní zobrazení které neobsahuje žádný cyklus , tj. $f(H) = H$ a $f^n(x) \neq x$ pro libovolný pár $(x, n) \in H \times \mathbb{N}$. Nechť $(H, *_f)$ je hypergrupa definovaná pomocí $x *_f y = R_f(x) \cup R_f(y)$ pro libovolný pár $x, y \in H$. Pak platí následující tvrzení:*

1. *Relace $R \subseteq H \times H$ je parametrizovatelná pomocí endomorfismů mono-unární algebry (H, f) právě tehdy, když je parametrizovatelná pomocí dobrých endomorfismů hypergrupy $(H, *_f)$.*
2. *Jakákoli relace $R \subseteq H \times H$ parametrizovatelná pomocí dobrých endomorfismů hypergrupy $(H, *_f)$ má substituční vlastnost (SP) on (H, f) a má substituční vlastnost prvního typu (SP-1) na $(H, *_f)$.*

Důkaz. (1) Na základě věty 7.1 z [51], s. 186 monoid endomorfismů mono-unární algebry (H, f) (tj. monoid všech zobrazení do sebe množiny H, což komutuje s f) korespondující s monoidem všech dobrých endomorfismů hypergrupy $(H, *_f)$. (Ve skutečnosti monoidy množinových transformací, které nesou odpovídající morfismus, jsou si rovny.) Z tohoto rovnostního tvrzení vyplývá bod (1) okamžitě.

(2) Nechť $x, y \in H$ je pár prvků takových, že xRy. Pak pro vhodný dobrý endomorfismus φ hypergrupy $((H, *_f)$ máme $\varphi \subseteq R$ a $\varphi(x) = y$. Protože podle [51], Věty 7.1 platí $\varphi \circ f = f \circ \varphi$, pak $f(y) = f(\varphi(x)) = \varphi(f(x))$, proto $f(x)Rf(y)$ a dostáváme poznatek, že relace R má SP vlastnost s ohledem na mono-unární algebru (H, f). Nyní předpokládejme $(a, b), (c, d) \in H \times H$, aRb, cRd. Pak zde jsou dobré endomorfismy $\varphi, \psi : (H, \cdot) \to (H, \cdot)$ takové, že $\varphi \subseteq R, \psi \subseteq R, \varphi(a) = b, \psi(c) = d$. Předpokládejme $x \in a \cdot c$. Pak pro nějaké $k \in \mathbb{N}_0 \backslash \{1\}$ buď $x = f^k(a)$, nebo $x = f^k(c)$. Protože podle Věty 7.1 [51]. $\varphi, \psi \in End(H, f)$, tj. $f\varphi = \varphi f$, $f\psi = \psi f$, máme buď

$$\varphi(x) = \varphi(f^k(a)) = f^k(\varphi(a)) = f^k(b) \in b \cdot d,$$

nebo

$$\psi(x) = \psi(f^k(c)) = f^k(\psi(c)) = f^k(d) \in b \cdot d,$$

takže zde existuje $y \in b \cdot d$, $y = \varphi(x)$ nebo $y = \psi(x)$ tak, že platí xRy. Jestliže $y \in b \cdot d$ pak — podobně jako výše pro nějaké $m \in \mathbb{N}_0 \backslash \{1\}$ buď $y = f^m(b)$, a nebo

$y = f^m(d)$ a pak buď

$$y = f^m(\varphi(a)) = \varphi(f^m(a)),\ \text{kde } f^m(a) \in a \cdot c,$$

nebo

$$y = f^m(\psi(c)) = \psi(f^m(c)),\ \text{kde } f^m(c) \in a \cdot c,$$

takže zde existuje $x \in a \cdot c$ tak, že platí xRy.Tudíž relace R má substituční vlastnost prvního typu (SP-1) na hypergrupě $(H, *_f)$. $\square$

Definice 10. *O binární relaci R na algebře (A, F) nebo respektive na uzávěrovém prostoru (A, C) řekneme že je homomorfní ve smyslu Grimeisena, krátce G-homomorfní, jestliže pro každou dvojici $a, b \in A$ s $\langle a, b\rangle \in R$ zde existuje $\varphi \in End(A, F)$ nebo respektive $\varphi \in End(A, C)$ tak, že platí $\varphi \subset R$ a $\varphi(a) = b$.*

Poznámka 3. Je zřejmé, že binární relace R na (A, F) (podobně pro uzávěrový prostor) je G-homomorfní právě tehdy, když zde existuje množina $\Phi \subset End(A, F)$ taková, že $R = \bigcup_{\varphi\in\Phi} \varphi$.

Definice 11. *O binární relaci R na algebře (A, F) řekneme, že je G-uzávěrově homomorfní, jestliže pro každou dvojici prvků $a, b \in A$ s $\langle a, b\rangle \in R$ zde existuje $\varphi \in End(A, C_F)$ tak, že platí $\varphi \subset R$ a $\varphi(a) = b$, tj. jestliže je G-homomorfní na (A, C_F).*

Je zřejmé, že homomorfní vztahy G-homomorfní a G-uzávěrově homomorfní jsou celodoménové. V následujícím textu jsou všechny uvažované vztahy celodoménové. Ve shodě s pojmem substituční vlastnost budeme také používat termíny: G-homomorfnost a G-uzávěrovou homomorfnost.

Tvrzení 3. *Každá G-homomorfní relace na libovolné množině stavů nějakého automatu je G-uzávěrově homomorfní.*

Poznámka 4. G-uzávěrová homomorfnost neimplikuje ani substituční vlastnost (SP), ani G-homomorfnost, (SP) neimplikuje ani G-homomorfnost ani G-uzávěrovou homomorfnost a G-homomorfnost obecně neimplikuje (SP). Vezměme tedy v úvahu 1-unární algebru $(\mathbb{N}, f)$,kde $f(n) = n + 1$ pro každé $n \in \mathbb{N}$. Definujme zobrazení φ, ψ množiny $\mathbb{N}$ do sebe tím, že položíme $\varphi(1) = 1$, $\varphi(n) = n - 1$ pro $n \geq 2$, $\psi(1) = \psi(2) = 1$ a $\psi(n) = n - 2$ pro $n \geq 3$. Položme $R = \varphi \cup \psi$. Pak R je evidentně G-homomorfní relace na $(\mathbb{N}, C_f)$, kde $C_f X = \bigcup_{0\leq k<\infty} f^k(X)$ pro každou podmnožinu $X \subset \mathbb{N}$, tj. R je G-uzávěrově homomorfní na $(\mathbb{N}, f)$. Platí $\langle 1, 1\rangle \in R$, ale $\langle f(1), f(1)\rangle = \langle 2, 2\rangle \notin R$, tím pádem R nemá (SP). Protože pro libovolné zobrazení $g : \mathbb{N} \to \mathbb{N}$ s vlastností $g \subset R$ zde platí $g(1) = g(2) = 1$, v důsledku toho $g \notin End(\mathbb{N}, f)$, a relace R není G-homomorfní.

Pro důkaz druhého tvrzení položme $G = \{a, b, c\}$ a označme zápisem $\mathfrak{G}$ komutativní grupoid (ve skutečnosti pologrupu) $(G, \circ)$, kde binární operace $\circ$ je dána pravidlem: Pro $x, y \in G$ položíme $x \circ y = b$ právě tehdy, když $x = y = a$ a $x \circ y = c$ v opačném případě. Nechť P je reflexní relace na G taková, že $\langle b, a\rangle \in P$ a $\langle c, b\rangle \in P$. Pak P má (SP) vzhledem k $\mathfrak{G}$, ale libovolné zobrazení $\varphi \subset P$ s $\varphi(b) = a$ není endomorfismus $\mathfrak{G}$ pro $\varphi(a) = a, \varphi(a \circ a) = a \neq b = \varphi(a) \circ \varphi(a)$. Dále (označujíc C jako korespondující algebraickou uzávěrovou operaci na $\mathfrak{G}$) máme $C\{a\} = G$, $C\{b\} = b, c$, $C\{c\} = \{c\}$, proto uvažované zobrazení φ není rovněž uzávěrový homomorfismus prostoru (G, C).

Nakonec uvažujme aditivní pologrupu kladných celých čísel $(\mathbb{N}, +)$ a reflexivní relaci $R = \left\{\langle m, n\rangle : m, n \in \mathbb{N}, \frac{n}{m} \in \{1, 2, 3\}\right\}$. Označme pomocí φ_1 zobrazení-identitu $id_{\mathbb{N}}$ a pomocí $\varphi_k : \mathbb{N} \to \mathbb{N}, k = 2, 3$ zobrazení definované pomocí $\varphi_k(n) = k \cdot n$ pro $k = 2, 3$ a každé $n \in \mathbb{N}$, máme $R = \bigcup\limits_{k=1}^{3} \varphi_k$ a zároveň $\varphi_k, k = 1, 2, 3$, jsou endomorfismy $(\mathbb{N}, +)$, takže R je G-homomorfní. Na druhé straně, $\langle 1, 2\rangle \in R$, $\langle 1, 3\rangle \in R$, ale $\langle 2, 5\rangle \notin R$. Je třeba poznamenat, že konečnost $\mathfrak{G}$ v první části výše uvedeného důkazu není pro daný protipříklad nezbytná. Může být nahrazen grupoidem $\mathfrak{G}_1 = (G_1, \circ)$ definovaným následovně: nechť $(G_1, \geq)$ je řetězec typu ω^* s největším prvkem s a pro $a, b \in G_1$ položíme $a \circ b = \min\{x : \max\{a, h\} < x\}$ jestliže $a \neq s \neq b$ a $a \circ b = s$ v opačném případě. Dále $\langle a, b\rangle \in R$, jestliže $a = b$ nebo a je předchůdce prvku b. Pokud se omezíme na unární algebry (s libovolným počtem unárních operací) dostaneme:

Tvrzení 4. *Každá G-homomorfní relace na libovolné unární algebře má substituční vlastnost (SP).*

Důkaz. Nechť R je G-homomorfní relace na unární algebře (A, F). Pro $a, b \in A$ takové, že $\langle a, b\rangle \in R$, zde existuje $\varphi \subset R, \varphi \in End(A, F)$ s $\varphi(a) = b$. Pak pro libovolnou operaci $f \in F$ máme $\varphi(f(a)) = f(\varphi(a)) = f(b)$, tj. $\langle f(a), f(b)\rangle \in R$ pro každou operaci $f \in F$. □

Věta 5. *Buď (A, f) 1-unární algebra, $\{(A_i, f_i); i \in I\}$ systém všech jejich komponentů, $I_0 = \{i \in I; cardA_i > 1\}, F = \{f^2, f^3\}$. Předpokládejme, že je splněna přesně jedna z těchto podmínek:*

1° *$i \in I$ implikuje buď $f_i^2 = f_i$* a nebo *$f_i^2 = id_{A_i}$,*

2° *$i \in I_0$ implikuje $A_i = A_i^{\infty 1}$,*

3° *$i \in I_0$ implikuje $A_i = A_i^{\infty 1} \cap A_i^0$ kde $(A_i^{\infty 1}, f_i)$ je obousměrný nekonečný řetězec a $A_i^0 \neq \emptyset$,*

4° *$i \in I_0$ implikuje buď (A_i, f_i) je obousměrný nekonečný řetězec, nebo $A_i = B_i \cap D_i$, kde $(D_i, f_i|D_i)$ je jednosměrný nekonečný řetězec s prvním prvkem $d, B_i = A_i^0$ a $f(x) = d$ pro každé $x \in B_i$.*

Pak každá G-uzávěrově homomorfní relace na 2-unární algebře (A, F) má (SP) vzhledem k (A, f) a (A, F).

Důkaz. Nechť $a, b \in A$ je pár prvků s vlastností $\langle a, b\rangle \in R$. Zde existuje endomorfismus g algebry (A, C_F) takový, že $g(a) = b$ a $\langle x, g(x)\rangle \in R$ pro každé $x \in A$. Poněvadž pro každé $a \in A$ platí $C_F\{a\} = \{f^k(a) : k = 0, 2, 3, ...\}$ a $C_F X = \bigcup_{x\in X} C_F F\{x\}$ pro $\emptyset \neq X \subset A$, máme dle Věty 3.2,4.1 z [52] $g \in End(A, f)$, takže $g(f(a)) = f(g(a)) = f(b)$. Proto $\langle f^k(a), f^k b\rangle \in R$ pro $k = 1, 2, 3,$. □

Z výše uvedené věty můžeme získat některé důsledky, např. následující:

Důsledek 1. Nechť $\mathfrak{U} = (A, \{f_1, f_2\})$ je 2-unární algebra taková, že f_1, f_2 jsou permutace množiny A a

1° $f_k^n h \neq h f_k^n$ pro každé $n \in \mathbb{N}$, libovolné konstantní zobrazení h množiny A do sebe a $k = 1, 2$,

2° $f_1 f_2 = f_2 f_1$,

3° $f_1^3 = f_2^2$

Pak každá G-uzávěrově homomorfní relace na algebře $\mathfrak{U}$ má substituční vlastnost (SP).

Důkaz. Buď $a \in A$ libovolný prvek a existuje zde $b \in A$ s vlastností $f_1(b) = a$. Položme $g(a) = f_2(b)$. Evidentně g je zobrazení množiny A do sebe. Pro každé $x \in A$ zde existuje $y \in A$ takové, že $f_1(y) = x$ a existuje $z \in A$ takové, že $f_1(z) = f_2(y)$. Pak $g^2(x) = g(g(x)) = g(f_2(y)) = f_2(z)$ a $f_1^2(x) = f_1^3(y) = f_2^2(y) = f_2 f_1(z) = f_1 f_2(z) = f_1(g^2(x))$, tj. $f_1^2 = f_1 g^2$. Poněvadž zobraze f_1 je vzájemně jednoznačné $f_1 = g^2$. Dále, $g^3(x) = g^2 g(x) = f_1 f_2(y) = f_2 f_1(y) = f_2 f(x)$, tj. $f_2 = g^3$. □

3.2 Modely grup a hypergrup umělých neuronů

3.2.1 Grupy a hypergrupy lineárních diferenciálních operátorů

Postup použitý při konstrukci algebraických struktur navazuje na práce šedesátých a sedmdesátých let minulého století, kdy začal výzkum diferenciálních rovnic pomocí algebraického a geometrického přístupu. Zejména Otakar Borůvka a jeho spolupracovníci rozvíjeli tuto cestu výzkumu.

Významný reprezentant uvedené školy František Neuman ve svém příspěvku [91] napsal: „Algebraické, topologické a geometrické nástroje spolu s metodami teorie dynamických systémů a funkcionálních rovnic umožňují řešit problémy týkající se globálních vlastností řešení v kontrastu k předchozím místním zkoumáním a izolovaným výsledkům.“ Tyto myšlenky jsou určitým motivačním faktorem dále popsaných konstrukcí.

Zvažujeme pro další postup lineární diferenciální operátory tvaru:

$$L_n = \sum_{k=0}^{n} p_k(x)D^k,$$

kde $D^k = \frac{d^k}{dx^k}$, $p_k(x)$ je spojitá funkce na nějakém otevřeném intervalu $J \subset \mathbb{R}$, $k = 0, 1, \dots, n-1$, $p_n(x) \equiv 1$,. Tedy $L_n(y) = 0$ je lineární homogenní obyčejná diferenciální rovnice tvaru:

$$y^{(n)}(x) + \sum_{k=0}^{n-1} p_k(x)y^{(k)}(x) = 0.$$

Uspořádanou grupou myslíme (jako obvykle) trojici $(G, \cdot, \leq)$, kde $(G, \cdot)$ je grupa a $\leq$ je reflexivní, symetrická and tranzitivní binární relace na množině G taková, že pro libovolnou trojici $x, y, z \in G$ s vlastností $x \leq y$ platí také $x \cdot z \leq y \cdot z$, $z \cdot x \leq z \cdot y$. Dále $[a)_{\leq} = \{x \in G; a \leq x\}$ je hlavní konec generovaný $a \in G$. K jakémukoli prvku $a \in G$ je přiřazena dvojice zobrazení $\lambda_a : G \to G, \rho_a : G \to G$, které se nazývají levá, respektive pravá translace[2], určená prvkem $a \in G$, tj. $\lambda_a(x) = a \cdot x$, $\rho_a(x) = x \cdot a$. (Samozřejmě v případě komutativní grupy $(G, \cdot)$ máme $\lambda_a \equiv \rho_a$).) Všimněme si, že grupa s uspořádáním $(G, \cdot, \leq)$ je uspořádaná grupa právě tehdy, když všechny její levé a pravé translace $\lambda_a, \rho_a, a \in G$ jsou uspořádání zachovávající, tj. izotonní zobrazení do sebe uspořádané množiny $(G, \leq)$.

Připomeňme v předchozím textu uvedenou Větu 2, známou jako „koncové lemma“, která je důležitá pro další konstrukci struktur umělých neuronů (dokázána v [55, 57], česká verze je dokázána v [51], s. 146, 147).

Nechť $(G, \cdot, \leq)$ *je uspořádaná grupa. Definujme hyperoperaci* $* : G \times G \to \mathcal{P}(G)^*$ *pomocí*

$$a * b = [a, b)_{\leq} (= \{x \in G; a \cdot b \leq x\})$$

pro každou dvojici prvků $a, b \in G$. *Pak* $(G, *)$ *je hypergrupa, která je komutativní právě tehdy, když grupa* $(G, \cdot)$ *je komutativní.*

Aplikace výše uvedeného lemmatu a mnoho nových výsledků jsou obsaženy v příspěvcích Michala Nováka. Pro příklad lze uvést alespoň tituly [94, 95, 96].

Pro popis následujících výsledků je užit podobný zápis, který byl publikovaný v [57]. $\mathbb{R}$ zde značí množinu všech reálných čísel, $J \subset \mathbb{R}$ je otevřený interval (ohraničený nebo neohraničený), $\mathbb{C}^k(J)$ je okruh (s ohledem na obvyklé sčítání a násobení funkcí) všech reálných funkcí se spojitými derivacemi až do řádu $k \geq 0$ včetně. Je

[2]Levé translace. Nechť $\mathcal{A}$ značí libovolnou grupu a a libovolný prvek v $\mathcal{A}$. Přiřadíme-li ke každému prvku $x \in \mathcal{A}$ prvek $ax \in \mathcal{A}$, obdržíme jisté zobrazení grupy $\mathcal{A}$ do sebe. Protože rovnice $ax = b$, v níž b značí libovolný prvek v $\mathcal{A}$, má jediné řešení $x \in \mathcal{A}$, je to prosté zobrazení grupy $\mathcal{A}$ na sebe, tj. permutace grupy $\mathcal{A}$. Tato permutace grupy $\mathcal{A}$ se nazývá levá translace určená prvkem a a označuje se ${}_a\mathbf{t}$. [18]

užit zápis $\mathbb{C}(J)$ místo $\mathbb{C}^0(J)$. Pro kladné celé číslo $n \geq 2$ je označena pomocí $\mathbb{A}_n$ množina všech lineárních homogenních diferenciálních rovnic n-tého řádu se spojitými reálnými koeficienty definovanými na J, tj.

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_0(x)y = 0,$$

(cf. [55, 57, 91, 92, 93]), kde $p_k \in \mathbb{C}(J)$, $k = 0, 1, \ldots, n-1$, $p_0(x) > 0$ pro libovolné $x \in J$ (to není zásadní omezení). Označme $L(p_0, \ldots, p_{n-1}) : \mathbb{C}^n(J) \to \mathbb{C}^n(J)$ výše uvedený lineární operátor definovaný pomocí

$$L(p_0, \cdots, p_{n-1})(y) = y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_0(x)y$$

a položme

$$\mathbb{LA}_n(J) = \{L(p_0, \cdots, p_{n-1}); p_k \in \mathbb{C}(J),\, p_0 > 0\}.$$

Dále $\mathbb{N}_0(n) = \{0, 1, \ldots, n-1\}$ a δ_{ij} je označení pro Kroneckerovo delta, $\overline{\delta_{ij}} = 1 - \delta_{ij}$. Pro libovolné $m \in \mathbb{N}_0(n)$ značíme jako $\mathbb{LA}_n(J)_m$ množinu všech lineárních diferenciálních operátorů $n-$tého řádu $L_0(p_0, \cdots, p_{n-1}) : \mathbb{C}^n(J) \to \mathbb{C}(J)$, kde $p_k \in \mathbb{C}(J)$ pro libovolné $k \in \mathbb{N}_0(n)$, $p_m \in \mathbb{C}_+(J)$, (tj. $p_m(x) > 0$ pro každé $x \in J$). Při použití vektorového zápisu $\vec{p}(x) = (p_0(x), \cdots, p_{n-1}(x))$, $x \in J$ můžeme psát pomocí skalárního součinu $L_n(\vec{p_0})y = y^{(n)} + (\vec{p}(x), (y, y', \cdots, y^{(n-1)}))$.

Nyní definujeme binární operaci „$\circ_m$“ a binární relaci $\leq_m$ na množině $\mathbb{LA}_n(J)_m$ takto:

Definice 12. *Pro libovolný pár $L(\vec{p})$, $L(\vec{q}) \in \mathbb{LA}_n(J)_m$, $\vec{p} = (p_0, \ldots, p_{n-1})$, $\vec{q} = (q_0, \ldots, q_{n-1})$ položíme $L(\vec{p}) \circ_m L(\vec{q}) = L(\vec{u})$, $\vec{u} = (u_0, \ldots, u_{n-1})$, kde*

$$u_k(x) = p_m(x)q_k(x) + (1 - \delta_{km})p_k(x),\ x \in J$$

a $L(\vec{p}) \leq L(\vec{q})$ kdykoli $p_k(x) \leq q_k(x)$, $k \in \mathbb{N}_0(n)$, $p_m(x) = q_m(x)$, $x \in J$.

Evidentně $(\mathbb{LA}_n(J)_m, \leq_m)$ je uspořádaná množina.

V článku [57] je uveden nástin důkazu následujícího lemmatu:

Lemma 1. *Trojice $(\mathbb{LA}_n(J)_m, \circ_m, \leq_m)$ je uspořádaná (nekomutativní) grupa.*

V následující kapitole je uvedena konstrukce grupy a hypergrupy umělých neuronů pomocí výše uvedeného přístupu.

3.2.2 Grupy a hypergrupy umělých neuronů

Neurony jsou často zmiňovány jako atomy neuronových výpočtů. Z těchto jednoduchých výpočetních jednotek jsou vybudovány všechny neuronové sítě. Výstup vypočítaný neuronem může být vyjádřen pomocí dvou funkcí $y = g(f(w, x))$. Podrobnosti výpočtu spočívají v několika krocích: V prvním kroku vstup do neuronu, $x := \{x_i\}$, je asociován s vahami[3] neuronových spojení, $w := \{w_i\}$, zapojením tzv. **propagační funkce** f. To lze považovat za výpočet aktivačního potenciálu z presynaptických činností. Z tohoto výsledku pak tzv. **aktivační funkce** g vypočítá výstup neuronu. Váhy, které napodobují synaptickou sílu, tvoří nastavitelné vnitřní parametry neuronu. Proces přizpůsobování vah (při trénování neuronové sítě) se nazývá učení.

Z biologického hlediska je vhodné použít integrační propagační funkci. A proto vhodnou volbou je použít vážený součet vstupu $f(w,x) = \sum_i w_i x_i$, což značí, že aktivační potenciál se rovná skalárnímu součinu vstupů a vah. Ve skutečnosti nejpopulárnější propagační funkce od úsvitu neuronových výpočtů je však často používána v mírně odlišné formě:

$$f(w,x) = \sum_i w_i x_i + \Theta, \tag{3.1}$$

Zvláštní váha Θ se nazývá bias. Aplikaci $\Theta(x) = 1$ pro $x > 0$ a $\Theta(x) = 0$ pro $x < 0$, jak ukazuje výše uvedená aktivační funkce, přináší slavný perceptron Rosenblatta. V tomto případě funguje funkce Θ jako práh.

Kromě (3.1) existují samozřejmě i další možné propagační funkce. Pokud je (3.1) doplněno o identitu jako aktivační funkce a skutečné domény mají skutečný lineární neuron získáme $\mathbf{y} = \sum_i w_i x_i + \Theta$. Tento skutečný lineární neuron lze považovat za příklad Cliffordova neuronu. Označuje $\mathfrak{C}_{p,q,r}$ jedinečnou univerzální Cliffordovu algebru odpovídající standardnímu kvadratickému prostoru $(\mathbb{R}^{p+q+r}, Q)$, $Q(x) = x^2$ a $\otimes_{p,q,r}$ geometrický součin algebry $\mathfrak{C}_{p,q,r}$ (srov. [20], kap. 2, část 2.1) můžeme říci, že Cliffordův neuron (CN) vypočítá následující funkci z $(\mathfrak{C}_{p,q,r})^n$ do $\mathfrak{C}_{p,q,r}$:

$$\mathbf{y} = \sum_{i=1}^{n} w_i \otimes_{p,q,r} x_i + \Theta.$$

([20], Definice 3.1).

Je třeba poznamenat, že hluboké neuronové sítě obsahují mnohonásobek nelineárních skrytých vrstev, což z nich dělá velmi expresivní modely, které se mohou naučit velmi komplikované vztahy mezi svými vstupy a výstupy.

Pro odhad parametrů sítě navrhl Hinton aj. [110] algoritmus wake-sleep. Jako modely slouží $q_i = \sigma\left(\sum_j s_i \Phi_{ij} + \Phi_{0j}\right)$ pro pozici „wake“a $p_j = \sigma\left(\sum_k s_k \Theta_{kj} + \Theta_{0j}\right)$ pro pozici „sleep“. Podobné pravděpodobnostní funkce se vyskytují v popisu modelu tzv. Omezených Boltzmannových strojů ([110].

Formální popis architektury obecné neuronové abstraktní pyramidy obsahuje formálně podobnou funkci uvedenou výše, danou matematickým modelem umělého neuronu. Základní prvek zpracování se skládá z projekčních jednotek P_{kl} a jedné

[3] Každému takovému spojení je přiřazena váhová hodnota $(w_1, w_2, \ldots, w_n)$, která určuje propustnost. [114]

výstupní jednotky. Aktivita $a^t_{ijkl} \in \mathbb{R}$ funkce buňky na pozici (i, j) pro pole funkcí k ve vrstvě l v čase t je vypočteno takto:

$$a^t_{ijkl} = \psi_{kl} \left(\sum_{p=1}^{P_{kl}} v^p_{kl} b^{l_p}_{ijkl} + v^0_{kl} \right).$$

Výstupní jednotka vypočítá vážený součet projekčních potenciálů $b^{tp}_{ijkl} \in \mathbb{R}$ s váhovými faktory popsanými pomocí $v^p_{kl} \in \mathbb{R}$. Do součtu je také přidána hodnota zkreslení (bias) v^0_{kl}, než je předána výstupní přenosovou funkcí ψ_{kl}.

Výpočet individuálních projekčních potenciálů popisuje:

$$b^{tp}_{ijkl} = \varphi_{kl} \left(\sum_{q=1}^{Q^p_{kl}} w^{pq}_{kl} a^{t'}{}_{i'j'k'l'} {}^{+w^{p0}_{kl}} \right).$$

Procesní prvek, který počítají funkční buňky, se skládá z projekčních jednotek P_{kl} a výstupní jednotky, která produkuje aktivitu a^t_{ijkl}. Výstupní jednotka vypočítá vážený součet potenciálů b^{tp}_{ijkl} jednotlivých projekcí a předá součet pomocí přenosové funkce ψ_{kl}. Každá projekční jednotka vypočítá vážený součet aktivit $a^{l'}_{i'j'k'l'}$ s váhovými faktory popsanými pomocí $w^{pq}_{kl} \in \mathbb{R}$. Počet příspěvků na projekci p je Q^p_{kl}. Kromě toho je před předáním součtu prostřednictvím projekční přenosové funkce ϕ^p_{kl} přidána hodnota zkreslení (bias) w^{p0}_{kl}.

Nyní si připomeňme známý matematický popis formálního neuronu [64]

Nechť $F : \mathbb{R} \to \mathbb{R}$ je obecná nelineární (nebo po částech lineární) funkce přenosu. Potom lze akci neuronu vyjádřit tímto modelem:

$$y(k) = F \left(\sum_{i=1}^{m} w_i(k) x_i(k) + b \right),$$

kde $x_i(k)$ je vstupní hodnota v diskrétním čase k, pro $i = 0, \ldots, m$, $w_i(k)$ je hodnota váhy v diskrétním čase, pro $i = 0, \ldots, m$, b je bias, $y_i(k)$ je výstupní hodnota v diskrétním čase k.

Všimněme si, že v některých velmi zvláštních případech může být přenosová funkce F také lineární. Přenosová funkce definuje vlastnosti umělého neuronu a může to být jakákoli matematická funkce. Obvykle se volí na základě problému, který umělý neuron (umělá neuronová síť) potřebuje vyřešit, a ve většině případů je převzata (jak je uvedeno výše) z následující množiny funkcí: skoková funkce, lineární funkce a nelineární (sigmoida) funkce.

Další výzkum je veden úvahou o určitém zobecnění klasických umělých neuronů zmíněných výše, spočívající v tom, že vstupy x_i a váhy w_i budou funkcemi argumentu t náležejícího do lineárně uspořádané (časové) množiny T s nejmenším prvkem 0. Jako množinu indexů používáme množinu $\mathbb{C}(J)$ všech spojitých funkcí definovaných v otevřeném intervalu $J \subset \mathbb{R}$. Označme tedy W množinu všech nezáporných funkcí

$w : T \to \mathbb{R}$ tvořících podpolookruh okruhu všech reálných funkcí jedné reálné proměnné $x : \mathbb{R} \to \mathbb{R}$. Označme pomocí $Ne(\vec{w}_r) = Ne(w_{r,1}, \ldots, w_{r,n})$ pro $r \in \mathbb{C}(J)$, $n \in \mathbb{N}$ zobrazení

$$y_r(t) = \sum_{k=1}^{n} w_{r,k}(t) x_{r,k}(t) + b_r,$$

které bude nazýváno umělým neuronem s biasem $b_r \in \mathbb{R}$. Jako $\mathbb{AN}(T)$ označujeme soubor všech takových umělých neuronů.

Neurony jsou obvykle označeny velkými písmeny X, Y nebo X_i, Y_i, nicméně používáme také symbolické značení $Ne(\vec{w})$, kde $\vec{w} = (w_1, \ldots, w_n)$ je vektor vah.

Předpokládáme pro zjednodušení že přenosové funkce (aktivační funkce) φ, σ (nebo f) jsou stejné pro všechny neurony ze souboru $\mathbb{AN}(T)$ nebo roli této funkce hraje identita $f(y) = y$.

Nyní podobně jako v případě souboru lineárních diferenciálních operátorů výše zkonstruujeme grupu a hypergrupu umělých neuronů. Označme jako δ_{ij} Kroneckerovo delta, $i, j \in \mathbb{N}$, tj. $\delta_{ii} = \delta_{jj} = 1$ a $\delta_{ij} = 0$, pro $i \neq j$.

Předpokládejme, že $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(T)$, $r, s \in \mathbb{C}(J)$, $\vec{w}_r = (w_{r1}, \ldots, w_{r,n})$, $\vec{w}_s = (w_{s1}, \ldots, w_{s,n})$,
$n \in \mathbb{N}$. Nechť $m \in \mathbb{N}$, $1 \leq m \leq n$ je takové celé číslo, že $w_{r,m} > 0$. Definujeme

$$Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s) = Ne(\vec{w}_u),$$

kde

$$\vec{w}_u = (w_{u,1}, \ldots, w_{u,n}) = (w_{u,1}(t), \ldots, w_{u,n}(t)),$$

$$\vec{w}_{u,k}(t) = w_{r,m}(t) w_{s,k}(t) + (1 - \delta_{m,k}) w_{r,k}(t), t \in T,$$

a samozřejmě, neuron $Ne(\vec{w}_u)$ je definován jako zobrazení $y_u(t) = \sum\limits_{k=1}^{n} w_k(t) x_k(t) + b_u$, $t \in T$, $b_u = b_r b_s$. Dále pro pár $Ne(\vec{w}_r)$, $Ne(\vec{w}_s)$ neuronů ze souboru $\mathbb{AN}(T)$ položíme $Ne(\vec{w}_r) \leq_m Ne(\vec{w}_s)$, $w_r = (w_{r,1}(t), \ldots, w_{r,n}(t))$, $w_s = (w_{s,1}(t), \ldots, w_{s,n}(t))$ jestliže $w_{r,k}(t) \leq w_{s,k}(t)$, $k \in \mathbb{N}$, $k \neq m$ a $w_{r,m}(t) = w_{s,m}(t)$, $t \in T$, se stejným biasem. Evidentně $(\mathbb{AN}(T), \leq_m)$ je uspořádaná množina.Vztah (kompatibility) binární operace „$\cdot_m$“ a uspořádání $\leq_m$ neuronů $\mathbb{AN}(T)$ je dáno tímto tvrzením analogickým s výše uvedeným.

Lemma 2. *Trojice* $(\mathbb{AN}(T), \cdot_m, \leq_m)$ *(algebraická struktura s uspořádáním) je uspořádaná nekomutativní grupa.*

Důkaz.
1. Operace „$\cdot_m$“ je evidentně asociativní.
2. Nechť $\vec{w}(t) = (w_1(t), \ldots, w_n(t))$, kde $w_k(t) = \delta_{km}$ pro libovolné $t \in T$. Pak neuron $Ne(\vec{w})$ určený pomocí $y(t) = \sum\limits_{k=1}^{n} w_k(t) x_k(t) + 1$ je neutrálním prvkem pologrupy $(\mathbb{AN}(T), \cdot_m)$, tj. jednotkou vzhledem k binární operaci „$\cdot_m$“.

3. Inverzní prvky: Definujme

$$\bar{w}_k(t) = (w_m(t))^{-1}(w_k(t)+1)\delta_{km} - w_k(t),$$

pro $t \in T$ a libovolné $k = 1, 2, \ldots, n$. Pak inverzní prvek k neuronu $Ne(\vec{w})$, s $y(t) = \sum_{k=1}^{n} w_k(t)x_k(t) + b$ v rámci pologrupy $(\mathbb{AN}(T), \cdot_m)$ je neuron $Ne(\vec{\bar{w}})$, určený pomocí $\bar{y}(t) = \sum_{k=1}^{n} \bar{w}_k(t)\bar{x}_k(t) + b^{-1}$.

4. Kompatibilita relace uspořádaní $\leq_m$ s binární operací „$\cdot_m$" (substituční vlastnost):

   Předpokládejme, že $Ne(\vec{w}_r)$, $Ne(\vec{w}_s)$, $Ne(\vec{w}_u) \in \mathbb{AN}(T)$ jsou takové neurony, že $Ne(\vec{w}_r) \leq_m Ne(\vec{w}_s)$, tj. $w_{r,m}(t) \equiv w_{s,m}(t)$, $w_{r,k}(t) \leq w_{s,k}(t)$ pro libovolný index $k \in \{1, 2, \ldots, n\}$, $k \neq m$, $t \in T$.

   Označme $Ne(\vec{w}_a) = Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_u)$, $Ne(\vec{w}_b) = Ne(\vec{w}_s) \cdot_m Ne(\vec{w}_u)$, kde $\vec{w}_a(t) = (w_{a,1}(t), \ldots, w_{a,n}(t))$, $\vec{w}_b(t) = (w_{b,1}(t), \ldots, w_{b,n}(t))$, $t \in T$.

   Pro libovolný index $k \in \{1, 2, \ldots, n\}$ máme

$$w_{a,k(t)} = w_{r,m}(t)w_{u,k}(t) + (1-\delta_{km})w_{r,k}(t),$$

$$w_{b,k}(t) = w_{s,m}(t)w_{u,k}(t) + (1-\delta_{km})w_{s,k}(t),\ t \in T,$$

   Poněvadž $w_{r,m}(t) \equiv w_{s,m}(t)$, $w_{r,k}(t) \leq w_{s,k}(t)$, $k \neq m$, $t \in T$, pro $k = m$ zde platí:

$$w_{a,m}(t) = w_{r,m}(t)w_{u,k}(t) = w_{s,m}(t)w_{u,m}(t),\ t \in T$$

   a pro $k \neq m$ máme

$$w_{a,k}(t) = w_{r,m}(t)w_{u,k}(t) + w_{r,k}(t) \leq w_{s,m}(t)w_{u,k}(t) + w_{s,k}(t) = w_{b,k}(t),\ t \in T,$$

   takže

$$Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_u) \leq_m Ne(\vec{w}_s) \cdot_m Ne(\vec{w}_u).$$

   Podobně získáváme platnost

$$Ne(\vec{w}_u) \cdot_m Ne(\vec{w}_r) \leq_m Ne(\vec{w}_u) \cdot_m Ne(\vec{w}_s). \qquad \square$$

Při označení

$$\mathbb{AN}_1(T)_m = \{Ne(\vec{w}); \vec{w} = (w_1, \ldots,\, w_n),\, w_k \in \mathbb{C}(T),\, k = 1, \ldots,\, n,\, w_m(t) \equiv 1\},$$

dostáváme následující tvrzení:

Tvrzení 5. *Nechť platí $T = [0, t_0) \subset \mathbb{R}$, $t_0 \in \mathbb{R} \cup \{\infty\}$. Potom pro jakékoli kladné celé číslo $n \in \mathbb{N}$, $n \geq 2$, a pro jakékoli celé číslo m takové, že $1 \leq m \leq n$, pologrupa $(\mathbb{AN}_1(T)_m, \cdot_m)$ je invariantní podgrupou grupy $(\mathbb{AN}(T)_m, \cdot_m)$.*

Důkaz. 1. Předpokládejme, že $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in (\mathbb{AN}_1(T)_m$ jsou libovolné neurony. Pak

$$\vec{w}_r = (w_{r,1}(t), \dots, w_{r,m}(t), \dots, w_{r,n}(t)),$$

$$\vec{w}_s = (w_{s,1}(t), \dots, w_{s,m}(t), \dots, w_{s,n}(t)),$$

kde $w_{r,m}(t) \equiv 1$, $w_{s,m}(t) \equiv 1$. Evidentně formální neuron $Ne(\vec{w})$,
kde $\vec{w}(t) = (w_1(t), \dots, w_n(t))$ s $w_k(t) = \delta_{km}$, $t \in T$, pracující prostřednictvím

$$y(t) = \sum_{k=1}^{n} w_k(t) x_k(t) + 1,$$

což je samozřejmě neutrální prvek grupy $(\mathbb{AN}(T)_m, \cdot_m)$ a také náleží k $\mathbb{AN}_1(T)_m$.
Označíme-li $Ne(\vec{w}_v) = Ne(\vec{w}_r) \cdot Ne(w_s)$, $\vec{w}_v = (w_{v,1}, \dots, w_{v,n})$ máme

$$w_{v,m}(t) = w_{r,m}(t)(w_{s,m}(t))^{-1} = 1$$

pro libovolné $t \in T$, takže $Ne(\vec{w}_v) \in \mathbb{AN}_1(T)_m$, v čehož důsledku získáme, že $(\mathbb{AN}_1(T)_m, \cdot_m)$ je podgrupa grupy $(\mathbb{AN}(T)_m, \cdot_m)$.

2. Nyní předpokládejme, že $Ne(\vec{w}_r) \in \mathbb{AN}(T)_m$, $Ne(\vec{w}_s) \in \mathbb{AN}_1(T)_m$, kde $\vec{w}_r(t)$, $\vec{w}_s(t)$ jsou vektory výše uvedené funkce. Při označení

$$Ne(\vec{w}_u) = Ne^{-1}(\vec{w}_r) \cdot_m Ne(\vec{w}_s) \cdot_m Ne(\vec{w}_r),$$

kde $\vec{w}_u(t) = \vec{w}(t) = (w_{u,1}(t), \dots, w_{u,n}(t))$, $t \in T$, pak platí $\vec{w}_{u,m}(t) = (\vec{w}_{r,m}(t))^{-1} \cdot_m \vec{w}_{s,m}(t) \cdot_m \vec{w}_{r,m}(t) = \vec{w}_{s,m}(t) = 1$ pro libovolné $t \in T$, takže $Ne(\vec{w}_u) \in \mathbb{AN}_1(T)_m$, což znamená, že

$$Ne^{-1}(\vec{w}_s) \cdot_m \mathbb{AN}_1(T)_m \cdot_m Ne(\vec{w}_s) \subset \mathbb{AN}_1(T)_m,$$

proto grupa $(\mathbb{AN}_1(T)_m, \cdot_m)$ je invariantní podgrupou grupy $(\mathbb{AN}(T)_m, \cdot_m)$.

□

Jestliže $m, n \in \mathbb{N}$, $1 \le m \le n-1$, pak určitý vztah mezi grupami $(\mathbb{AN}_r(T)_m, \cdot_m)$, $(\mathbb{LA}(T)_{m+1}, \circ_{m+1})$ je obsažen v následujícím tvrzení:

Tvrzení 6. *Nechť $t_0 \in \mathbb{R}$, $t_0 > 0$, $T = [0, t_0) \subset \mathbb{R}$ a $m, n \in \mathbb{N}$ jsou celá čísla, taková, že $1 \le m \le n-1$. Definujme zobrazení $F : \mathbb{AN}_n(T)_m \to \mathbb{LA}_n(T)_{m+1}$ tímto pravidlem: pro libovolný neuron $Ne(\vec{w}_r \in \mathbb{AN}_n(T)_m$, kde $\vec{w}_r = (w_{r,1}(t), \dots, w_{r,n}(t)) \in [\mathbb{C}(T)]^n$ položíme $F(Ne(\vec{w}_r)) = L(w_{r,1}, \dots, w_{r,n}) \in \mathbb{LA}_n(T)_{m+1}$ s akcí:*

$$L(w_{r,1}, \dots, w_{r,n}) y(t) = \frac{d^n y(t)}{dt^n} + \sum_{k=1}^{n} w_{r,k}(t) \frac{d^{k-1}(t)}{dt^{k-1}},\ y \in \mathbb{C}^n(T).$$

Pak zobrazení $F : \mathbb{AN}_n(T)_m \to \mathbb{LA}_n(T)_{m+1}$ je homomorfismus grupy $(\mathbb{AN}_n(T)_m, \cdot_m)$ do grupy $(\mathbb{LA}_n(T)_{m+1}, \circ_{m+1})$.

Důkaz. Uvažujme $Ne(\vec{w}_r),\ Ne(\vec{w}_s) \in \mathbb{AN}_n(T)_m$ a označme $F(Ne(\vec{w}_r)) = L(w_{r,1}, \ldots, w_{r,n})$, $F(Ne(\vec{w}_s)) = L(w_{s,1}, \ldots, w_{s,n})$. Označme $Ne(\vec{w}_u) = Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s)$. Zde platí

$$F(Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s)) = F(Ne(\vec{w}_u)) = L(w_{u,1}, \ldots, w_{u,n}),$$

kde

$$L(w_{u,1}, \ldots, w_{u,n})y(t) = y^{(n)}(t) + \sum_{k=1}^{n} w_{u,k}(t)y^{(k-1)}(t).$$

Platí $w_{u,k}(t) = w_{r,m+1}(t)w_{s,k}(t) + w_{r,k}(t)$, $k \neq m$, a $w_{u,m+1}(t) = w_{r,m+1}(t)w_{s,m+1}(t)$. Pak je $L(w_{u,1}, \ldots, w_{u,n}) = L(w_{r,1}, \ldots, w_{r,n}) \cdot_m L(w_{s,1}, \ldots, w_{s,n}) = F(Ne(\vec{w}_r)) \cdot_m F(Ne(\vec{w}_s))$. Neutrální prvek $Ne(\vec{w}) \in \mathbb{AN}_n(T)_m$ je tak zobrazen na neutrální prvek grupy $(\mathbb{L}_n\mathbb{A}(T)_{m+1}, \cdot_{m+1})$, tedy zobrazení $F : (\mathbb{AN}_n(T)_m, \cdot_m) \to (\mathbb{L}_n\mathbb{A}(T)_{m+1}, \circ_{m+1})$ je grupový homomorfismus. $\square$

Nyní pomocí konstrukce popsané ve výše uvedeném Lemmatu 2 získáme konečnou transpoziční hypergrupu (nazývanou také nekomutativní spojovací prostor). Označme $\mathbb{P}(\mathbb{AN}(T)_m)^*$ potenční množinu $\mathbb{AN}(T)_m$ sestávající ze všech neprázdných podmnožin této množiny a definujme binární hyperoperaci

$$*_m : \mathbb{AN}(T)_m \times \mathbb{AN}(T)_m \to \mathbb{P}(\mathbb{AN}(T)_m)^*$$

podle pravidla

$$Ne(\vec{w}_r) *_m Ne(\vec{w}_s) = \{Ne(\vec{w}_u);\ Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s) \leq_m Ne(\vec{w}_u)\}$$

pro všechny páry $Ne(\vec{w}_r), Ne(\vec{w}_s) \in \mathbb{AN}(T)_m$. Podrobněji, pokud $\vec{w}(u) = (w_{u,1}, \ldots, w_{u,n})$, $\vec{w}(r) = (w_{r,1}, \ldots, w_{r,n})$, $\vec{w}(s) = (w_{s,1}, \ldots, w_{s,n})$, pak $w_{r,m}(t)w_{s,m}(t) = w_{u,m}(t)$, $w_{r,m}(t)w_{s,k}(t) + w_{r,k}(t) \leq w_{u,k}(t)$, jestliže $k \neq m$, $t \in T$. Potom máme, že $(\mathbb{AN}(T)_m, *_m)$ je nekomutativní hypergrupa. Výše definovaná invariantní (nazývaná také normální) podgrupa $(\mathbb{AN}_1(T)_m, \cdot_m)$ grupy $(\mathbb{AN}(T)_m, \cdot_m)$ je nesená množina podhypergrupy hypergrupy $(\mathbb{AN}(T)_m, *_m)$ a má určité významné vlastnosti.

Pomocí zobecnění metod ze článku [57] získáme po prozkoumání konstruovaných struktur tento výsledek:

Věta 6. *Buď* $T = [0, t_0) \subset \mathbb{R}$, $t_0 \in \mathbb{R} \cup \{\infty\}$. *Pak pro každé přirozené číslo* $n \in \mathbb{N}$, $n \geq 2$, *a pro jakékoli celé číslo* m *takové, že* $1 \leq m \leq n$, *hypergrupa* $(\mathbb{AN}(T)_m, *_m)$, *kde*

$$\mathbb{AN}(T)_m = \{Ne(\vec{w}_r);\ \vec{w}_r = (w_{r,1}(t), \ldots, w_{r,n}(t)) \in [\mathbb{C}(T)]^n,\ w_{r,m}(t) > 0,\ t \in T\},$$

je transpoziční hypergrupa (tj. nekomutativní spojovací prostor) taková,
že $(\mathbb{AN}(T)_m, *_m)$ *je její subhypergrupa, která je*

- *invertibilní (tj.* $Ne(\vec{w}_r)/Ne(\vec{w}_s) \cap \mathbb{AN}_1(T)_m \neq \emptyset$ *implikuje* $Ne(\vec{w}_s)/Ne(\vec{w}_r) \cap \mathbb{AN}_1(T)_m \neq \emptyset$ *a* $Ne(\vec{w}_r) \setminus Ne(\vec{w}_s) \cap \mathbb{AN}_1(T)_m \neq \emptyset$ *implikuje* $Ne(\vec{w}_s) \setminus Ne(\vec{w}_r) \cap \mathbb{AN}_1(T)_m \neq \emptyset$ *pro každou dvojici neuronů* $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}_1(T)_m$,
- *uzavřená (tj.* $Ne(\vec{w}_r)/Ne(\vec{w}_s) \subset \mathbb{AN}_1(T)_m$, $Ne(\vec{w}_r) \setminus Ne(\vec{w}_s) \subset \mathbb{AN}_1(T)_m$ *pro každou dvojici* $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}_1(T)_m$,
- *reflexivní (tj.* $Ne(\vec{w}_r) \setminus \mathbb{AN}_1(T)_m = \mathbb{AN}_1(T)_m/Ne(\vec{w}_r)$ *pro libovolný neuron* $Ne(\vec{w}_r) \in \mathbb{AN}(T)_m$ *a*
- *normální (tj.* $Ne(\vec{w}_r) * \mathbb{AN}_1(T)_m = \mathbb{AN}_1(T)_m * Ne(\vec{w}_r)$ *pro libovolný neuron* $Ne(\vec{w}_r) \in \mathbb{AN}(T)_m$.

Poznámka 5. Určité zobecnění formálního (umělého) neuronu lze získat expresí lineárního diferenciálního operátoru n-tého řádu. Připomeňme výraz formálního neuronu s vnitřním potenciálem $y_{-in} = \sum_{k=1}^{n} w_k(t)x_k(t)$, kde $\vec{x}(t) = (x_1(t), \ldots, x_n(t))$ je vektor vstupů, $\vec{w}(t) = (w_1(t), \ldots, w_n(t))$ je vektor vah. Užitím biasu b uvažovaného neuronu a přenosové funkce σ můžeme výstup vyjádřit jako

$$y(t) = \sigma\left(\sum_{k=1}^{n} w_k(t)x_k(t) + b\right).$$

Nyní uvažujme kmenovou funkci $u : J \to \mathbb{R}$, kde $J \subseteq \mathbb{R}$ je otevřený interval; vstupy jsou odvozeny od funkce $u \in \mathbb{C}^n(J)$ následovně: Vstupy $x_1(t) = u(t), x_2 = \frac{du(t)}{dt}, \ldots, x_n(t) = \frac{d^{n-1}(t)}{dt^{n-1}}$, $n \in \mathbb{N}$. Dále bias $b = b_0 \frac{d^n u(t)}{dt^n}$. Jako váhy používáme spojitou funkci $w_k : J \to \mathbb{R}$, $k = 1, \ldots, n-1$.

Pak vzorec

$$y(t) = \sigma\left(\sum_{k=1}^{n} w_k(t) \frac{d^{k-1}u(t)}{dt^{k-1}} + b_0 \frac{d^n u(t)}{dt^n}\right)$$

je popis akce neuronu Dn který bude nazýván formálním (umělým) diferenciálním neuronem. Tento přístup umožňuje využít prostorů řešení odpovídajících lineárních diferenciálních rovnic.

3.3 Model iterovaných časově proměnných neuronů

V lineárních modelech je výstupem modelu lineární funkce vstupních proměnných. Existuje několik způsobů, jak mohou být tyto funkce zobecněny. Jednou z metod je použití monotónní nelineární aktivační funkce $g(a)$, která zpracovává lineární kombinaci prvků ve tvaru:

$$f(x, w) = g\left(w_0 + \sum_{i=1}^{d} w_i x_i\right).$$

Jednou pozoruhodnou volbou pro aktivační funkci je logistická sigmoidní funkce, která zobrazuje interval $(-\infty, \infty)$ na interval $(0, 1)$:

$$g(a) = \frac{1}{1 - \exp(a)}.$$

Funkce $f(x, w)$ se nazývá diskriminační funkce, protože odděluje třídy navzájem. Ve statistické literatuře jsou často uváděny modely s logistickou sigmoidní aktivační funkcí jako logistická diskriminace.

V příspěvku [38] je definována „sendvičová" operace „$\circ$" pro jakýkoli prvek a pologrupy $(S, \cdot)$ na stejné množině S. Operace „$\circ$" je definována pravidlem $x \circ y = x \cdot a \cdot y$, $(x, y \in S)$. V rámci této operace je množina S opět pologrupou; autor příspěvku [38] označuje tuto pologrupu (S, a) a nazývá ji variantou S. Varianty pologrup jsou zkoumány v článcích [70, 38], kde lze nalézt další literaturu, včetně článků věnovaných variantám pologrup binárních vztahů zpracovaných Karem Chase. Uvádíme zde odstavec 5 [38] týkající se částečného uspořádání pro regulární pologrupy. Připomeňme, že pologrupa S je považována za regulární, pokud je jakýkoli prvek $x \in S$ regulární, tj. existují $y \in S$ takové, že $xyx = x$. Připomeňme, že pro pravidelnou pologrupu S částečné Nambooripadovo uspořádání $\leq$ na S může být definované jako

$$x \leq y \text{ právě tehdy, když } x \in y \cdot S \cdot y \text{ a } x = x \cdot y' \cdot x$$

pro některé $y' \in V(y)$. Množina $V(y)$ je definována jako

$$V(y) = \{u \in S; y \cdot u \cdot y = y \text{ a } u \cdot y \cdot u = u\}.$$

Další důležitý poznatek se nachází v článku [38], kde autor dokázal Větu 5.1: *Buď S regulární pologrupa a nechť relace ρ je definovaná na S jako*

$$x\rho y \text{ právě tehdy, když existuje } a \in S \text{ takové, že } x, y \in E((S, a)),$$

(podmnožina všech idempotentních prvků varianty (S, a)), jež splňuje podmínku $x \circ y = y \circ x = x$ z (S, a). Pak ρ shoduje se s částečným Nambooripadovým uspořádáním na S a tak je sám o sobě částečným uspořádáním na S.

S ohledem na právě zmíněnou větu od Hickeye můžeme binární strukturu $\mathbb{AN}(T)$ modifikovanou jako variantu pologrupy umělých časově proměnných neuronů s nosičem $\mathbb{AN}(T)$. Je známo, že monoid veškerých lineárních transformací vektorového prostoru je regulární (ve smyslu konceptu regularity J. von Neumanna), takže můžeme vytvořit pravidelnou podpologrupu grupy $\mathbb{AN}(T)$, protože transformace určené neurony s nulovými biasy jsou ve skutečnosti lineární. Tato pologrupa je označena jako $\mathbb{SAN}(T)$, pologrupa může být vybavena Nambooripadovým uspořádáním $\leq$, a tak můžeme vytvořit množinu komutativní rozsáhlé hypergrupy na množině $\mathbb{SAN}(T)$, kde hyperproduktem libovolných dvou neuronů je sjednocení horních konců (tj. kuželů) generovaných každým z těchto neuronů. Je zřejmé, že univerzální binární operace na jakékoli variantě pologrupy je zvláštním případem Vougiouklisovy binární P-hyperoperace [116, 117].

Připomeňme, že hypergroupoid je dvojice $(G, \cdot)$, kde $\cdot : G \times G \to \mathcal{P}(G)$ je binární hyperoperace. Pokud je tato hyperoperace asociativní, tj. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ pro jakoukoli trojici prvků $a, b, c \in G$ (kde $a \cdot (b \cdot c) = \bigcup \{a \cdot x | x \in b \cdot c\}$), pak hypergroupoid $(G, \cdot)$ se nazývá polohypergrupa. Polohypergrupa $(G, \cdot)$, která splňuje reprodukční axiom (tj. $a \cdot G = G = G \cdot a$ pro jakékoli $a \in G$) se nazývá hypergrupa - srov. [66, 23, 24]. Nyní předpokládejme, že $(G, \cdot)$, $(H, \cdot)$ jsou polohypegrupy. Zobrazení $f : G \to H$ je považováno za homomorfismus (resp. dobrý homomorfismus) polohypergrupy $(G, \cdot)$ do polohypergrupy $(H, \cdot)$, pokud pro libovolný pár prvků $x, y \in G$ máme $f(x \cdot y) \subseteq f(x) \cdot f(y)$ (resp. $f(x \cdot y) = f(x) \cdot f(y)$).

Definujme nyní koncept P-hypergroupy podle Thomase Vougiouklise [116, 117]:

Definice 13. *Nechť $(G, \cdot)$ je pologrupa a $P \subset G$, $P \neq \emptyset$. Hyperoperace $*^P : G \times G \to \mathcal{P}(G)$ definovaná jako $[x, y] \to xPy$, i.e. $x * y = xPy$ pro libovolný pár $[x, y] \in P \times P$ je označována jako P-hyperoperace v G. Pokud*

$$x *^P (y *^P z) = xPyPz = (x *^P y) *^P z$$

platí pro každou trojici $x, y, z \in G$, pak je P-hyperoperace asociativní. Pokud je uspokojen i axiom reprodukce, je hypergrupoid $(G, *^P)$ považován za P-hypergrupu.

Je zřejmé, že pokud $(G, \cdot)$ je grupa, pak také $(G, *^P)$ je P-hypergrupa. Pokud je množina P singleton[4], pak je P-operace $*^P$ obvyklou jedno-hodnotovou operací.

Definice 14. *O podmnožině $H \subset G$ řekneme, že je a sub-P-hypergrupa hypergrupy $(G, *^P)$ jestliže platí $P \subset H \subset G$ a $(H, *^P)$ je hypergrupa.*

Tvrzení 7. *([115] str.16). Nechť $(G_1, \cdot)$, $(G_2, \cdot)$ jsou dvě grupy, $f \in Hom(G_1, G_2)$ a $P \subset G_1$. Pak je homomorfismus f dobrým homomorfismem mezi P-hypergrupami $(G_1, *^P)$ a $(G_2, *^P)$.*

Více o diskutovaných tématech viz [70, 78, 115, 116, 117].

Označme tedy $S \subseteq \mathbb{C}(T)$ libovolnou neprázdnou podmnožinu a nechť platí

$$P = \{Ne(\vec{w}_u(t)); u \in S\} \subseteq \mathbb{AN}(T).$$

Pak definujeme

$$Ne(\vec{w}_p(t)) * Ne(\vec{w}_q(t)) = Ne(\vec{w}_p(t)) \cdot_m P \cdot_m Ne(\vec{w}_q(t)) =$$

[4] V matematice, singleton, také známý jako jednotková množina, je množina s přesně jedním prvkem. Například množina $\{0\}$ je singleton obsahující prvek 0. Termín je také používán pro 1-tici (posloupnost s jedním členem) [112].

$$= \{Ne(\vec{w}_p(t)) \cdot_m Ne(\vec{w}_u(t)) \cdot_m Ne(\vec{w}_q(t)); u \in S\}$$

pro libovolný pár neuronů $Ne(\vec{w}_p(t)), Ne(\vec{w}_q(t)) \in \mathbb{AN}(T)$, zde získáme P-hypergrupu umělých časově proměnlivých neuronů. V případě, že množina S je singleton, tj. P je jednoprvková podmnožina $\mathbb{AN}(T)$, získaná struktura je variantou $\mathbb{AN}(T)$. Všimněme si, že libovolný homomorfismus $f \in EndG$ pro grupu $(G, \cdot)$ indukuje dobrý homomorfismus P-hypergrup $(G, *^P)$, $(G, *^{f(P)})$ a jakýkoli automorfismus vytváří izomorfismus mezi výše uvedenými P-hypergrupami.

Hypergrupy uvažované v tomto případě jsou hypergrupy ve smyslu Martyho [82].

3.3.1 Struktury ve vícevrstvé perceptronové síti (MLP)

V monografii [76] bylo navrženo několik různých architektur neuronových sítí pod dohledem, založených na modelu umělého perceptronu. Podle zmiňovaného autora (Timo Koskela) jsou nejoblíbenější sítí vícevrstvé perceptronové sítě (MLP) a sítě radiální funkce základů (RBF)[5]. V části odstavce 3.3 zdroje [76] (Supervised Learning) je krátce zhodnocena populární vícevrstvá síť Perceptron a jsou představeny síťové architektury založené na strukturách zpoždění a opakujících se připojeních.

Vícevrstevná perceptronová síť (MLP) je nelineární model sestávající z množství neuronů (jednotek) uspořádaných do více vrstev, které vytvářejí zobrazení $y = f(x, w)$ mezi vstupem x a výstupem y, upravený váhami w. Složitost MLP sítě lze změnit z téměř lineárního modelu na vysoce nelineární model změnou počtu vrstev, počtu jednotek v každé vrstvě a hodnot vah. Typická jediná skrytá vrstva architektury MLP sítě s K výstupy vede k modelu $f_k(x, w), k = 1, \ldots, K$ s váhami w. Funkční formu modifikovanou časově proměnnými vstupy a vahami pro $t \in T$ popisujeme takto:

$$\begin{aligned} f_k(x(t), w(t)) = w_0(t) + w_1(t)\Big(w_{1,0}(t) + \sum_{i=1}^{d} w_{1,i}(t)x_i(t)\Big) + \\ + w_2(t)\Big(w_{2,0}(t) + \sum_{i=1}^{d} w_{2,i}(t)x_i(t)\Big) + \cdots + \\ + w_{m-1}(t)\Big(w_{m-1,0}(t) + \sum_{i=1}^{d} w_{m-1,i}(t)x_i(t)\Big) + \\ + w_m(t)\Big(w_{m,0}(t) + \sum_{i=1}^{d} w_{m,i}(t)x_i(t)\Big). \end{aligned}$$

[5]Neuronové sítě typu RBF (Radial Basis Functions) jsou třívrstvé neuronové sítě. Základní vlastností těchto sítí je že aktivační funkce je typu RBF. RBF funkce jsou takové které jsou určeny svým středem a pro argumenty se stejnou vzdáleností od tohoto středu dávají stejné funkční hodnoty [76].

MLP je často považován za obecný semi-parametrický model, což znamená, že efektivní počet parametrů může být menší než počet dostupných parametrů. Efektivní počet parametrů určuje složitost modelu. Pro malé váhy je síťové zobrazení téměř lineární a má nízkou efektivní složitost. Je to proto, že centrální oblast sigmoidální aktivační funkce může být aproximována pomocí a lineární funkce.

Při označení logistické aktivační funkce g a pokud indexy i, j odpovídají výstupním a skrytým jednotkám, model $f(x, w)$ s váhami w dává tvar:

$$
\begin{aligned}
f(x(t), w(t)) = g\Big(w_0(t)\Big) + w_1(t) g\Big(w_{1,0}(t) + \sum_{i=1}^{d} w_{1,i}(t) x_i(t)\Big) + \\
+ w_2(t) g\Big(w_{2,0}(t) + \sum_{i=1}^{d} w_{2,i}(t) x_i(t)\Big) + \cdots + \\
+ w_{m-1}(t) g\Big(w_{m-1,0}(t) + \sum_{i=1}^{d} w_{m-1,i}(t) x_i(t)\Big) + \\
+ w_m(t) g\Big(w_{m,0}(t) + \sum_{i=1}^{d} w_{m,i}(t) x_i(t)\Big),
\end{aligned}
$$

kde časový parametr t patří do časové škály T.

Jako určitý model používající iteraci umělých neuronů lze použít konstrukci kaskády, tj. akci aditivní grupy $(\mathbb{Z}, +)$ všech celých čísel na grupě $\mathbb{AN}(T)$. Tato struktura je motivována konstrukcí uvedenou ve zdroji [54] s. 48.

Obvykle je regularizace potřebná, aby poskytovala dobrou generalizační schopnost pro nepoužité vzory během tréninku. Tradičně složitost MLP byla kontrolována předčasným zastavením nebo metodou rozpadu vah [29].

Dopředné vícevrstvé sítě jsou architektury, kde jsou neurony spojeny do vrstev a spojení mezi vrstvami jde pouze jedním směrem, ze vstupní vrstvy do výstupní vrstvy. Mezi neurony ve stejné vrstvě není žádné spojení. Mezi vstupní a výstupní vrstvou může také být jedna nebo několik skrytých vrstev.

Buď $(\mathbb{Z}, +)$ aditivní grupa všech celých čísel. Nechť $Ne(\vec{w}_s(t) \in \mathbb{AN}(T)$ je libovolný, ale pevně daný umělý neuron s výstupní funkcí $y_s(t) = \sum_{k=1}^{n} w_{s,k}(t) x_{s,k}(t) + b_s$. Označme jako $\lambda_s : \mathbb{AN}(T) \to \mathbb{AN}(T)$ levou translaci v grupě časových proměnných neuronů určenou prvkem $Ne(\vec{w}_s(t))$, tj.

$$
\lambda_s(Ne(\vec{w}_p(t))) = Ne(\vec{w}_s(t)) \cdot_m Ne(\vec{w}_p(t))
$$

pro libovolný neuron $Ne(\vec{w}_p(t)) \in \mathbb{AN}(T)$. Dále označme pomocí λ_s^r r-tou iteraci λ_s pro $r \in \mathbb{Z}$. Definujme projekci $\pi_s : \mathbb{AN}(T) \times \mathbb{Z} \to \mathbb{AN}(T)$ pomocí

$$
\pi_s(Ne(\vec{w}_p(t)), r) = \lambda_s^r(Ne(\vec{w}_p(t)).
$$

Je snadné vidět, že dostáváme obvyklou (diskrétní) transformační grupu, tj. akci $(\mathbb{Z}, +)$ (jako fázová grupa) na grupu $\mathbb{AN}(T)$, takže jsou splněny následující dva požadavky:

1. $\pi_s(Ne(\vec{w}_p(t)), 0) = Ne(\vec{w}_p(t))$ pro libovolný umělý neuron $Ne(\vec{w}_p(t)) \in \mathbb{AN}(T)$,
2. $\pi_s(Ne(\vec{w}_p(t)), r+u) = \pi_s(\pi_s(Ne(\vec{w}_p(t)), r), u)$ pro všechna celá čísla $r, u \in \mathbb{Z}$ a libovolný umělý neuron $Ne(\vec{w}_p(t))$. Poznamenejme, že právě získaná struktura se v rámci teorie dynamického systému nazývá kaskáda.

Na fázové množině, tj. nosiči grupy $\mathbb{AN}(T)$, definujeme binární hyperoperaci (určenou tzv. stavovou hypergrupou automatu). Takže pro libovolný pár neuronů $Ne(\vec{w}_p(t)), Ne(\vec{w}_q(t))$ definujme

$$Ne(\vec{w}_p(t)) * Ne(\vec{w}_q(t)) = \pi_s(Ne(\vec{w}_p(t)), \mathbb{Z}) \cup \pi_s(Ne(\vec{w}_q(t)), \mathbb{Z}) =$$

$$\{\lambda_s^a(Ne(\vec{w}_p(t)); a \in \mathbb{Z}\} \cup \{\lambda_s^b(Ne(\vec{w}_q(t)); b \in \mathbb{Z}\}.$$

Potom máme $* : \mathbb{AN}(T) \times \mathbb{AN}(T) \to \mathcal{P}(\mathbb{AN}(T))$ je komutativní binární hyperoperace a protože $Ne(\vec{w}_p(t)), Ne(\vec{w}_q(t)) \in Ne(\vec{w}_p(t)) * Ne(\vec{w}_q(t))$, dostaneme poznatek, že hypergroupoid $(\mathbb{AN}(T), *)$ je komutativní extenzivní hypergrupa [54, 66, 23, 24, 95, 116]. Pomocí jeho vlastností lze charakterizovat určité vlastnosti kaskády $(\mathbb{AN}(T), \mathbb{Z}, \pi_s)$. Hypergrupu $(\mathbb{AN}(T), *)$ lze nazvat fázovou hypergrupou dané kaskády.

Mezi elementárními pojmy teorie dynamických systémů je koncept invariantních podmnožin fázové množiny kaskády $(X, \mathbb{Z}, \pi_s)$ a koncept kritického bodu. Podmnožina M fázové množiny X kaskády $(X, \mathbb{Z}, \pi_s)$ se nazývá invariantní, kdykoli $\pi(x, r) \in M$, pro všechna $x \in M$ a všechna $r \in \mathbb{Z}$. Kritickým bodem kaskády je invariantní singleton. Je zřejmé, že podmnožina M neuronů, tj. $M \subseteq \mathbb{AN}(T)$, je invariantní v kaskádě $(\mathbb{AN}(T), \mathbb{Z}, \pi_s)$, kdykoli se jedná o nosnou množinu podhypergrupy hypergrupy $(\mathbb{AN}(T), *)$, tj. M je uzavřena s ohledem na hyperoperaci $*$, což znamená $M * M = \bigcup_{a,b \in M} a * b \subseteq M$. Kromě toho je invariantní také sjednocení nebo průnik libovolného neprázdného systému $\mathcal{M} \subseteq \mathbb{AN}(T)$.

3.4 Řešitelnost grup a hypergrup umělých neuronů

Pro prezentaci následujících výsledků výzkumu algebraických struktur na množině umělých neuronů je užitečné připomenout některé základní pojmy z teorie řad grup. Více podrobností viz např. [11, 60, 66, 30, 31, 35, 68]. Je-li G grupa,H její normální podgrupa, tj. normální dělitel, píšeme $H \lhd G$[6]. Posloupnost podgrup $H_i\,(i = 0, 1, \ldots, n)$ grupy G takových, že

$$1 = H_0 \lhd H_1 \lhd \cdots \lhd H_{n-1} \lhd H_n = G \tag{3.2}$$

je považován za subnormální řadu grupy G, odpovídající faktorové grupy H_i/H_{i-1} $(i = 1, 2, \ldots, n)$ jsou nazývány faktory řady $(\mathcal{H})$.

[6]Pojmy normální podgrupa, normální dělitel a invariantní podgrupa splývají v teorii grup.

Definice 15. *([11, 57, 60, 35]). Jestliže G je grupa, pak o její subnormální řadě*

$$1 = H_0 \lhd H_1 \lhd \cdots \lhd H_{n-1} \lhd H_n = G$$

řekneme, že je řešitelná, jestliže všechny faktory $H_i/H_{i-1}(i = 1, 2, ..., n)$ jsou komutativní, tedy jsou abelovskými grupami. Pokud grupa G má alespoň jednu řešitelnou subnormální řadu, pak se G nazývá řešitelná grupa.

Poznámka 6. Je zřejmé, že libovolná abelovská grupa je řešitelná, nicméně existují řešitelné grupy, které nejsou abelovské. V literatuře lze nalézt dobře známé výsledky týkající se řešitelných grup. Zmiňme některé z nich. Zaprvé je to slavná Feitova–Thompsonova věta aneb věta o lichém řádu, která říká, že každá konečná grupa lichého řádu je řešitelná. Dokázali to Walter Feit a John Griggs Thompson (1962, 1963) [30, 31]. Dále každá konečná grupa řádu $|G| \leq 100$ kromě 60 je řešitelná. (Za tyto informace a další podrobnosti jsou oba autoři podle svých slov velmi zavázáni profesoru Václavu Havlovi.) Libovolná p-grupa G, tj. $|G| = p^n$ pro nějaké prvočíslo p je řešitelná. Je-li $K \lhd G$, pak grupa G je řešitelná právě tehdy, když K a G/K jsou řešitelné. Grupa je řešitelná právě tehdy, když obsahuje subnormální řadu, jejíž všechny faktory jsou řešitelné. Zejména přímým produktem konečného počtu řešitelných grup je opět řešitelná grupa. Ovšem žádná konečná grupa G stupně 5 a více není řešitelná. Další důležité informace lze nalézt v literatuře [11, 60, 66, 68, 30, 31, 35]. Připomeňme, že podgrupa G' grupy G generovaná množinou komutátorů

$$[a, b] = a^{-1}b^{-1}ab$$

všech párů $[a, b] \in G \times G$ se nazývá komutant grupy G. Komutant G'' grupy G' se nazývá druhý komutant grupy G. Obvyklý zápis je $G' = G^{(1)}$, $G'' = G^{(2)}$, atd. Obecně $G^{(n+1)} = (G^{(n)})'$. Komutant $G^{(n)}$ je také nazýván jako n-tý derivát grupy G a řetěz komutantů grupy G,

$$G = G^{(0)} \supset G^{(1)} \supset G^{(2)} \supset \cdots$$

se také nazývá odvozený řetězec grupy G.

Věta 7. *(Věta 1. [57], [103], Věta 5., s.117) Buď G grupa. Následující podmínky jsou ekvivalentní:*

(i) Grupa G je řešitelná.

(ii) Existuje kladné celé číslo n takové $G^{(n)} = 1$.

(iii) Existuje subnormální řada grupy G, která je řešitelná.

Nyní můžeme použít reverzní přístup. Z důvodů, které jsou zřejmé z dalšího textu, je použito následující označení:

Příklad Představme konkrétní příklad. Pojmenujme strukturu $[\mathbb{C}(J)]^3 = \mathbb{C}(J) \times \mathbb{C}(J) \times \mathbb{C}(J)$ jako funkcionální krychli, pak označme grupu $([\mathbb{C}(J)]^3_2, \circ_2)$, s prvky $\vec{p}(x), \vec{q}(x) \in [\mathbb{C}(J)]^3$, $\vec{p}(x) = (p_0(x), p_1(x), p_2(x))$, $\vec{q}(x) = (q_0(x), q_1(x), q_2(x))$, poté definujeme

$$[\mathbb{C}(J)_2]^3 = \{\vec{p}(x) = (p_0(x), p_1(x), p_2(x)); p_k \in \mathbb{C}(J), p_2(x) \neq 0, x \in J\},$$

a binární operaci $\circ_2 : [\mathbb{C}(J)_2]^3 \times [\mathbb{C}(J)_2]^3 \to [\mathbb{C}(J)_2]^3$ definujme takto:

$$\begin{aligned}\vec{p}(x) \circ_2 \vec{q}(x) &= (p_0(x), p_1(x), p_2(x)) \circ_2 (q_0(x), q_1(x), q_2(x)) \\ &= \vec{w}(p_2(x)q_0(x) + p_0(x), p_2(x)q_1(x) + p_1(x), p_2(x)q_2(x)),\end{aligned}$$

$$\text{pak } (\vec{p}(x))^{-1} = (p_0(x), p_1(x), p_2(x))^{-1} = \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right) = \vec{v}(x).$$

Tyto formulace s analogií operátorů mají dostatečně obecnou podobu, která ukazuje cestu k vyřešení výše zmíněného problému. Nicméně řešení tohoto problému (tj. otázka řešitelnost grupy $(\mathbb{C}(J)_n, \circ_m)$) se zdá být prozatím otevřená.

Na základě výše uvedených úvah získáme poznatek, že $[\mathbb{C}(J)_2, \circ_2]^3$ je nekomutativní grupa s neutrálním prvkem $\vec{e}(x) = (0, 0, 1)$ (kde $\vec{e}(x) \circ_2 \vec{p}(x) = \vec{p}(x)$ pro libovolný prvek $\vec{p}(x) \in [\mathbb{C}(J)_2, \circ_2]^3$. Inverzní prvek k $\vec{p}(x) = (p_0(x), p_1(x), p_2(x))$ je uveden výše. Je zřejmé, že

$$\vec{p}(x) \circ_2 (\vec{p}(x))^{-1} = (p_0(x), p_1(x), p_2(x)) \circ_2 \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right)$$

$$= \left(-\frac{p_2(x)p_0(x)}{p_2(x)} + p_0(x), -\frac{p_2(x)p_1(x)}{p_2(x)} + p_1(x), \frac{p_2(x)}{p_2(x)}\right) = (0, 0, 1) = \vec{e}(x).$$

Podobně získáme neutrální prvek ze součinu výše uvedených vektorů z funkcionální krychle v opačném pořadí.

Označme

$$[\mathbb{C}_{2C}(J)_2]^3 = \{(p_0(x), p_1(x), r); p_0, p_1 \in \mathbb{C}(J)\}, r \in R, r \neq 0$$

a

$$[\mathbb{C}_{21}(J)_2]^3 = \{(p_0(x), p_1(x), 1); p_0, p_1 \in \mathbb{C}(J)\}.$$

Pak můžeme vyslovit větu:

Věta 8. *Nechť $J \subseteq \mathbb{R}$ je otevřený interval. Grupoidy $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ jsou podgrupy grupy $([\mathbb{C}(J)_2]^3, \circ_2)$ - první z nich je komutativní - a máme*

$$([\mathbb{C}_{21}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}_{2C}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}(J)_2]^3, \circ_2).$$

Důkaz. Evidentně

$$(0,0,1) \in [\mathbb{C}_{21}(J)_2]^3 \subseteq [\mathbb{C}_{2C}(J)_2]^3.$$

Pokud $(p_0(x), p_1(x), r)$, $(q_0(x), q_1(x), s) \in [\mathbb{C}_{2C}(J)_2, \circ_2]^3$ (s $r, s \in \mathbb{R}, r \neq 0 \neq s$) jsou vektory spojitých funkcí, pak máme:

$$(p_0(x), p_1(x), r) \circ_2 (q_0(x), q_1(x), s)^{-1} = (p_0(x), p_1(x), r) \circ_2 \left(-\frac{q_0(x)}{s}, -\frac{q_1(x)}{s}, \frac{1}{s}\right)$$

$$= \left(p_0(x) - \frac{r}{s} q_0(x), p_1(x) - \frac{r}{s} q_1(x), \frac{r}{s}\right) \in [\mathbb{C}_{2C}(J)_2]^3.$$

Podobně

$$(p_0(x), p_1(x), 1) \circ_2 (q_0(x), q_1(x), 1) \in [\mathbb{C}_{2C}(J)_2]^3.$$

Takže $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ jsou podgrupy grupy $([\mathbb{C}(J)_2]^3, \circ_2)$. Dále,

$$(p_0(x), p_1(x), 1) \circ_2 (q_0(x), q_1(x), 1) = (p_0(x) + q_0(x), p_1(x) + q_1(x), 1)$$

$$= (q_0(x), q_1(x), 1) \circ_2 (p_0(x), p_1(x), 1),$$

pro libovolné vektory spojitých funkcí $(p_0(x), p_1(x), 1)$, $(q_0(x), q_1(x), 1) \in [\mathbb{C}_{21}(J)_2]^3$, v důsledku čehož grupa $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$ je komutativní.

Nyní pro libovolné vektory $(p_0(x), p_1(x), p_2(x)) \in [\mathbb{C}(J)_2]^3$ a $(q_0(x), q_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$ získáme:

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 (q_0(x), q_1(x), r) \circ_2 (p_0(x), p_1(x), p_2(x))$$

$$= \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right) \circ_2 (rp_0(x) + q_0(x), rp_1(x) + q_1(x), rp_2(x))$$

$$= (\varphi_0(x), \varphi_1(x), r)\,, \text{ kde } \varphi_k(x) = r(p_k(x) + q_k(x)),\ k = 0, 1.$$

Protože $(\varphi_0(x), \varphi_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$, platí

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 [\mathbb{C}_{2C}(J)_2]^3 \circ_2 (p_0(x), p_1(x), p_2(x)) \subseteq [\mathbb{C}_{2C}(J)_2]^3,$$

tudíž $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ je normální podgrupa grupy $([\mathbb{C}(J)_2]^3, \circ_2)$.

Podobným způsobem ověříme, že

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 [\mathbb{C}_{21}(J)_2]^3 \circ_2 (p_0(x), p_1(x), p_2(x)) \subseteq [\mathbb{C}_{21}(J)_2]^3$$

pro libovolný vektor $(p_0(x), p_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$. Z toho důvodu

$$([\mathbb{C}_{21}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}_{2C}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}(J)_2]^3, \circ_2).$$

$\square$

Věta 9. *Nechť $J \subseteq \mathbb{R}$ je otevřený interval. Grupa $([\mathbb{C}(J)_2]^3, \circ_2)$ všech vektorů spojitých funkcí je řešitelná.*

Důkaz. Označme $G = ([\mathbb{C}(J)_2]^3, \circ_2)$. První derivace G', tj. první komutant grupy G je jeho podgrupa generovaná všemi komutátory tvaru

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 (q_0(x), q_1(x), q_2(x))^{-1}$$

$$\circ_2 (p_0(x), p_1(x), p_2(x)) \circ_2 (q_0(x), q_1(x), q_2(x)),$$

kde $(p_0(x), p_1(x), p_2(x))$, $(q_0(x), q_1(x), q_2(x)) \in [\mathbb{C}(J)_2]^3$. Vzhledem k výše uvedenému komutátoru

$$\begin{aligned}
&(\vec{p}(x))^{-1} \circ_2 (\vec{q}(x))^{-1} \circ_2 \vec{p}(x) \circ_2 \vec{q}(x) \\
&= \left(-\frac{p_0(x)q_2(x) + p_0(x)}{p_2(x)q_2(x)}, -\frac{p_1(x)q_2(x) + q_1(x)}{p_2(x)q_2(x)}, \frac{1}{p_2(x)q_2(x)}\right) \\
&\circ_2 (p_2(x)q_0(x) + p_0(x), p_2(x)q_1(x) + p_1(x), p_2(x)q_2(x)) \\
&= \left(\frac{q_0(x)(p_2(x) - 1) + p_0(x)(1 - q_2(x))}{p_2(x)q_2(x)}, \frac{q_1(x)(p_2(x) - 1) + p_1(x)(1 - q_2(x))}{p_2(x)q_2(x)}, 1\right)
\end{aligned}$$

náleží do grupy $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, platí,že G' je podgrupa poslední výše uvedené grupy. Uvažujme libovolný pár vektorů $(u_0(x), u_1(x), 1)$, $(v_0(x), v_1(x), 1) \in ([\mathbb{C}_{21}(J)_2]^3, \circ_2)$. Druhá derivace G, tj. druhý komutant G'' grupy $G = ([\mathbb{C}(J)_2]^3, \circ_2)$ je jeho podgrupa generovaná množinou všech komutátorů

$$(u_0(x), u_1(x), 1)^{-1} \circ_2 (v_0(x), v_1(x), 1)^{-1} \circ_2 (u_0(x), u_1(x), 1) \circ_2 (v_0(x), v_1(x), 1)$$

$$= (-u_0(x), -u_1(x), 1) \circ_2 (-v_0(x), -v_1(x), 1) \circ_2 (u_0(x) + v_0(x), u_1(x) + v_1(x), 1)$$

$$= (-(u_0(x)+v_0(x)), -(u_1(x)+v_1(x)), 1) \circ_2 (u_0(x)+v_0(x), u_1(x)+v_1(x), 1) = (0, 0, 1)$$

proto $G'' = \{(0, 0, 1)\}$ což je triviální podgrupa (tvořená pouze jednotkou) grupy $G = ([\mathbb{C}(J)_2]^3, \circ_2)$. V důsledku čehož

$$\{(0, 0, 1)\} = G'' \subset G' \subset G^{(0)} = G,$$

proto grupa $([\mathbb{C}(J)_2]^3, \circ_2)$ je řešitelná.

Povšimněme si, že abychom rozhodli, zda grupa G' je ekvivalentní s $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$ nebo ne, vyžaduje další ověření. To spočívá v otázce, zda libovolnou spojitou funkci $f \in \mathbb{C}$ lze vyjádřit ve formě koeficientů operátoru

$$L^{-1}(\vec{p}(x)) \circ_2 L^{-1}(\vec{q}(x)) \circ_2 L(\vec{p}(x)) \circ_2 L(\vec{q}(x)).$$

□

Věta 10. *Výše definovaná grupa $(\mathbb{AN}(T), \bullet_m)$ umělých neuronů je řešitelná.*

Důkaz. Označme $G = (\mathbb{AN}(T), \bullet_m)$. První derivace G' grupy G, což znamená první komutant grupy G, je její podgrupa generovaná všemi komutátory tvaru

$$Ne^{-1}(\vec{w}_r) \bullet_m Ne^{-1}(\vec{w}_s) \bullet_m Ne(\vec{w}_r) \bullet_m Ne(\vec{w}_s),$$

kde $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(T)$. Protože uvedený komutátor má podobu

$$Ne\left(-\frac{w_{r,1}}{w_{r,m}}, \ldots, \frac{1}{w_{r,m}}, \ldots, -\frac{w_{r,n}}{w_{r,m}}\right) \bullet_m Ne\left(-\frac{w_{s,1}}{w_{s,m}}, \ldots, \frac{1}{w_{s,m}}, \ldots, -\frac{w_{s,n}}{w_{s,m}}\right) \bullet_m$$
$$Ne\left(w_{r,m}w_{s,1} + w_{r,1}, \ldots, w_{r,m}w_{s,m}, \ldots, w_{r,m}w_{s,n} + w_{r,n}\right) =$$
$$Ne\left(-\frac{w_{s,m}w_{r,1} + w_{s,1}}{w_{r,m}w_{s,m}}, \ldots, \frac{1}{w_{r,m}w_{s,m}}, \ldots, -\frac{w_{s,m}w_{r,n} + w_{s,n}}{w_{r,m}w_{s,m}}\right) \bullet_m$$
$$Ne\left(w_{r,m}w_{s,1} + w_{r,1}, \ldots, w_{r,m}w_{s,m}, \ldots, w_{r,m}w_{s,n} + w_{r,n}\right) =$$
$$Ne\left(\frac{(w_{r,m} - 1)w_{s,1} + (1 - w_{s,m})w_{r,1}}{w_{r,m}w_{s,m}}, \ldots, 1, \ldots, \frac{(w_{r,m} - 1)w_{s,n} + (1 - w_{s,m})w_{r,n}}{w_{r,m}w_{s,m}}\right)$$
$$\in \mathbb{AN}_1(T)_m,$$

získáme poznatek, že podgrupou grupy $(\mathbb{AN}(T), \bullet_m)$ generované všemi komutátory výše popsaného tvaru je právě grupa $(\mathbb{AN}_1(T)_m, \bullet_m)$. Předpokládejme, že k je kladné celé číslo, $1 \leq k \leq n$, $k \neq m$ a $f \in \mathbb{C}(T)$ je libovolná funkce a poté volbou $w_{r,m}(t) = 1$, $w_{s,m}(t) = t^2 + 1$ a $w_{r,k}(t) = -\frac{f(t)}{t^2}(t^2 + 1)$ zjistíme, že

$$\frac{1}{w_{r,m}(t)w_{s,m}(t)}\left((w_{r,m}(t) - 1)w_{s,k}(t) + (1 - w_{s,m}(t)w_{r,k}(t)\right) =$$
$$\frac{-t^2 w_{r,k}(t)}{t^2 + 1} = f(t),\ t \in T.$$

V popsaném tvaru tedy může být vyjádřena libovolná kontinuální váha z výše uvedeného neuronu. Proto $G' = (\mathbb{AN}_i(T)_m, \bullet_m)$. Dále uvažujme libovolný pár neuronů

$$Ne(w_{u,1}, \ldots, 1, \ldots, w_{u,n}),\ Ne(w_{v,1}, \ldots, 1, \ldots, w_{v,n}) \in \mathbb{AN}_1(T)_m.$$

Na základě úvah, které jsou obdobné výše uvedeným, máme:

$$Ne^{-1}(w_{u,1}, \ldots, 1, \ldots, w_{u,n}) \bullet_m Ne^{-1}(w_{v,1}, \ldots, 1, \ldots, w_{v,n}) \bullet_m$$
$$Ne(w_{u,1}, \ldots, 1, \ldots, w_{u,n}) \bullet_m Ne(w_{v,1}, \ldots, 1, \ldots, w_{v,n}) =$$
$$Ne(-w_{u,1}, \ldots, 1, \ldots, -w_{u,n}) \bullet_m Ne(-w_{v,1}, \ldots, 1, \ldots, -w_{v,n}) \bullet_m$$
$$Ne(w_{u,1} + w_{v,1}, \ldots, 1, \ldots, w_{u,n} + w_{v,n}) =$$
$$Ne(-(w_{u,1}+w_{v,1}), \ldots, 1, \ldots, -(w_{u,n}+w_{v,n})) \bullet_m Ne(w_{u,1}+w_{v,1}, \ldots, 1, \ldots, w_{u,n}+w_{v,n})$$
$$= Ne(0, \ldots, 0, 1, 0, \ldots, 0),$$

tedy druhá derivace G'' je triviální grupa, která obsahuje pouze identitu.

Proto jsme získali jako výsledek

$$\{Ne(0, \ldots, 0, 1, 0, \ldots, 0)\} = G'' \subseteq G' = G^{(0)} = G = (\mathbb{AN}(T), \bullet_m),$$

z toho důvodu je grupa $(\mathbb{AN}(T), \bullet_m)$ řešitelná. $\square$

3.5 Řada polohypergrup časově proměnných umělých neuronů a souvisejících hyperstruktur

Další konstrukce navazuje na výše definované pojmy hypergrupoid, polohypergrupa a hypegrupa, definované v předchozím popisu algebraických struktur.

Hypergrupy, které jsou popsané v následující kapitole, jsou struktury ve smyslu F. Martyho [49, 82]. Připomeňme také užití lemmatu (nazývaného také jako koncové lemma -srovnej [61, 94, 95, 96, 97]). Jako (kvazi-)uspořádanou pologrupu chápeme trojici $(S, \cdot, \leq)$, kde $(S, \cdot)$ je pologrupa, $(S, \leq)$ je (kvazi-)uspořádaná množina, tj. množina S mající reflexivní a tranzitivní binární relaci „$\leq$“ a pro všechny trojice prvků $a, b, c \in S$ platí implikace $a \leq b \Rightarrow a \cdot c \leq b \cdot c$, $c \cdot a \leq c \cdot b$.

Pro další výzkum množin umělých neuronů použijeme již uvedenou Větu 2.

Nechť $(S, \cdot, \leq)$ je (kvazi-)uspořádaná pologrupa. Definujme binární hyperoperaci

$$* : S \times S \to \mathcal{P}^*(S) \text{ jako } a * b = \{x \in S; a \cdot b \leq x\}.$$

Pak $(S, *)$ je polohypergrupa. Navíc, jestliže pologrupa $(S, \cdot)$ je komutativní, pak je polohypergrupa $(S, *)$ také komutativní a když $(S, \cdot, \leq)$ je (kvazi-)uspořádaná grupa, pak je polohypergrupa $(S, *)$ také hypergrupou.

Nyní budeme konstruovat řadu grup a hypergrup umělých neuronů pomocí určité analogie s řadou grup diferenciálních operátorů popsaných v [61].

Označme pomocí $\mathbb{LA}_n(J)$ (pro otevřený interval $J \subseteq \mathbb{R}$) množinu všech lineárních diferenciálních operátorů $L(p_{n-1}, \ldots, p_0)$, $p_0 \neq 0$, $p_k \in \mathbb{C}^n(J)$, tj. okruh všech spojitých funkcí definovaných na intervalu J, chovající se jako

$$L(p_{n-1}, \ldots, p_0)y(x) = y^n(x) + \sum_{k=0}^{n-1} p_k(x)y^k(x),\ y \in \mathbb{C}^n(J)$$

a obsahující binární operaci

$$L(q_{n-1}, \ldots, q_0) \circ L(p_{n-1}, \ldots, p_0) = L(q_0 p_{n-1} + q_{n-1}, \ldots, q_0 p_1 + q_1, q_0 p_0).$$

Nyní označme jako $\overline{\mathbb{LA}}_n(J)$ množinu všech operátorů $\overline{L}(q_n, \ldots, q_0)$, $q_0 \neq 0$, $q_k \in \mathbb{C}(J)$ v akci jako

$$\overline{L}(q_n, \ldots, q_0)y(x) = \sum_{k=0}^{n} q_k(x)y^{(k)}(x),\ q_0 \neq 0,\ q_k \in \mathbb{C}(J)$$

s podobně definovanými binárními operacemi, takže $\mathbb{LA}_n(J)$, $\overline{\mathbb{LA}}_n(J)$ jsou nekomutativní grupy. Definujme zobrazení $F_n : \mathbb{LA}_n(J) \to \mathbb{LA}_{n-1}(J)$ pomocí

$$F_n(L(p_{n-1}, \ldots, p_0)) = L(p_{n-2}, \ldots, p_0)$$

a $\phi_n : \mathbb{LA}(J) \to \overline{\mathbb{LA}}_{n-1}(J)$ pomocí

$$\phi_n(L(p_{n-1}, \ldots, p_0)) = \overline{L}(p_{n-2}, \ldots, p_0).$$

Tak lze snadno ověřit, že obě zobrazení F_n a ϕ_n, pro libovolné $n \in \mathbb{N}$, jsou grupové homomorfismy.

$$\begin{tikzcd}
\overline{\mathbb{LA}}_n(J) \arrow[r, "\overline{id}_{n,n+1}"] & \overline{\mathbb{LA}}_{n+1}(J) \arrow[r, "\overline{id}_{n+1,n+2}"] & \overline{\mathbb{LA}}_{n+2}(J) \arrow[r, "\overline{id}_{n+2,n+3}"] & \overline{\mathbb{LA}}_{n+3}(J) \arrow[r, "\overline{id}_{n+3,n+4}"] & \cdots \\
\mathbb{LA}_n(J) \arrow[u, "id_n"] & \mathbb{LA}_{n+1}(J) \arrow[l, "F_{n+1}"] \arrow[u, "id_{n+1}"] \arrow[ul, "\phi_{n+1}"] & \mathbb{LA}_{n+2}(J) \arrow[l, "F_{n+2}"] \arrow[u, "id_{n+2}"] \arrow[ul, "\phi_{n+2}"] & \mathbb{LA}_{n+3}(J) \arrow[l, "F_{n+3}"] \arrow[u, "id_{n+3}"] \arrow[ul, "\phi_{n+3}"] & \cdots \arrow[l, "F_{n+4}"] \arrow[ul, "\phi_{n+4}"]
\end{tikzcd}$$

Nyní zvažme grupy časově proměnných neuronů $(\mathbb{AN}(T)_m, \cdot_m)$ z Tvrzení 7 a výše definovaný homomorfismus grupy $(\mathbb{AN}_n(T)_m, \cdot_m)$ do grupy $(\mathbb{LA}_n(T)_{m+1}, \circ_{m+1})$. Potom můžeme změnit diagram následujícím způsobem:

$$\begin{tikzcd}
\overline{\mathbb{AN}}_n(T)_m \arrow[r, "\overline{id^*}_{n,n+1}"] & \overline{\mathbb{AN}}_{n+1}(T)_m \arrow[r, "\overline{id^*}_{n+1,n+2}"] & \overline{\mathbb{AN}}_{n+2}(T)_m \arrow[r, "\overline{id^*}_{n+2,n+3}"] & \overline{\mathbb{AN}}_{n+3}(T)_m \arrow[r, "\overline{id^*}_{n+3,n+4}"] & \cdots \\
\mathbb{AN}_n(T)_m \arrow[u, "id^*_n"] & \mathbb{AN}_{n+1}(T)_m \arrow[l, "F^*_{n+1}"] \arrow[u, "id^*_{n+1}"] \arrow[ul, "\phi^*_{n+1}"] & \mathbb{AN}_{n+2}(T)_m \arrow[l, "F^*_{n+2}"] \arrow[u, "id^*_{n+2}"] \arrow[ul, "\phi^*_{n+2}"] & \mathbb{AN}_{n+3}(T)_m \arrow[l, "F^*_{n+3}"] \arrow[u, "id_{n+3}"] \arrow[ul, "\phi^*_{n+3}"] & \cdots \arrow[l, "F^*_{n+4}"] \arrow[ul, "\phi^*_{n+4}"]
\end{tikzcd}$$

Pomocí *Koncového lemmatu* a výsledků teorie lineárních operátorů můžeme popsat také zobrazení morfismů v posloupnostech grup lineárních diferenciálních operátorů:

$$\mathbb{LA}_n(J) \xleftarrow[F_{n+1}]{} \mathbb{LA}_{n+1}(J) \xleftarrow[F_{n+2}]{} \mathbb{LA}_{n+2}(J) \xleftarrow[F_{n+3}]{} \mathbb{LA}_{n+3}(J) \xleftarrow[F_{n+4}]{} \cdots$$

tak jako analogicky v posloupnostech grup časově proměnných neuronů:

$$\mathbb{AN}_n(T)_m \xleftarrow[F^*_{n+1}]{} \mathbb{AN}_{n+1}(T)_m \xleftarrow[F^*_{n+2}]{} \mathbb{AN}_{n+2}(T)_m \xleftarrow[F^*_{n+3}]{} \mathbb{AN}_{n+3}(T)_m \xleftarrow[F^*_{n+4}]{} \cdots \tag{3.3}$$

Věta 11. *Nechť $T = [0, t_0) \subset \mathbb{R}$, $t_0 \in \mathbb{R} \cup \{\infty\}$, $n \in \mathbb{N}$ takové, že $n \geq 2$, $m \in \mathbb{N}$ takže $m \leq n$. Nechť $(\mathbb{HAN}_n(T)_m, *_m)$ je hypergrupa získaná z grupy $(\mathbb{AN}_n(T)_m, \circ_m)$ pomocí Tvrzení 6. Předpokládejme, že $F_n : (\mathbb{AN}_n(T)_m, \circ_m) \to (\mathbb{AN}_{n-1}(T)_m, \circ_m)$ jsou výše definované surjektivní grupy homomorfismů. Pak $F_n : (\mathbb{HAN}_n(T)_m, *_m) \to (\mathbb{HAN}_{n-1}(J)_m, *_m)$ jsou surjektivní homomorfismy hypergrup.*

Poznámka 7. Druhá posloupnost (3.3) může být tedy bijektivně zobrazena na posloupnost hypergrup

$$\mathbb{HAN}_n(T)_m \xleftarrow[F^*_{n+1}]{} \mathbb{HAN}_{n+1}(T)_m \xleftarrow[F^*_{n+2}]{} \mathbb{HAN}_{n+2}(T)_m \xleftarrow[F^*_{n+3}]{} \mathbb{HAN}_{n+3}(T)_m \xleftarrow[F^*_{n+4}]{} \cdots$$

s vazebnými surjektivními homomorfismy F_n. Proto je bijektivní zobrazení výše uvedených posloupností funktoriální.

Nyní přejděme ke konceptu automatu, který byl vyvinut jako matematická interpretace systémů reálného života, které pracují na diskrétním časovém měřítku. Pomocí binární operace zřetězení řetězců vstupních symbolů získáváme automaty se vstupními abecedami disponujícími strukturou pologrupy nebo grupy. Berouc v úvahu hlavně strukturu danou přechodovou funkcí a zanedbávaje výstupní funkce s výstupními množinami získáme velmi užitečné zobecnění konceptu automatu zvaného kvazi-automat [8, 19, 61, 24].

Představme si tedy koncept automatů jako akci časově proměnlivých neuronů. Navíc nechť systém (A, S, δ) sestává z množiny stavů neprázdných časově proměnných neuronů $A \subseteq \mathbb{AN}(T)_m$, libovolné pologrupy jejich vstupů S a nechť zobrazení $\delta : A \times S \to A$ splňuje následující podmínku:

$$\delta(\delta(a, r), s) = \delta(a, r \cdot s)$$

pro libovolné $a \in A$ a $r, s \in S$. Podmínku lze chápat jako analogii konceptu kvazi-automatu, jako zobecnění automatu Mealyho typu. Výše uvedená podmínka se někdy nazývá Podmínka smíšené asociativity (MAC - Mixed Associativity Condition).

Definice 16. *Nechť A je neprázdná množina, $(H, \cdot)$ polohypergrupa a $\delta : A \times H \to A$ zobrazení splňující podmínku:*

$$\delta(\delta(s, a), b) \in \delta(s, a \cdot b) \tag{3.4}$$

pro libovolnou trojici $(s, a, b) \in A \times H \times H$, kde $\delta(s, a \cdot b) = \{\delta(s, x); x \in a \cdot b\}$. Trojice (A, H, δ) se nazývá kvazi-multiautomat se stavovou množinou A a vstupní polohypegrupou $(H, \cdot)$. Zobrazení $\delta : A \times H \to A$ se nazývá přechodová funkce (nebo funkce následujícího stavu) kvazi-multiautomatu (A, H, δ). Podmínka (3.4) se nazývá Generalizovaná podmínka smíšené asociativity (nebo-li GMAC - Generalized Mixed Associativity Condition).

Právě definované struktury se také nazývají akcemi polohypergrup $(H, \cdot)$ na množinách A (nazývaných stavové množiny).

Neuron $Ne(\vec{w})$ se chová jak je popsáno výše:

$$y(t) = \sum_{i=1}^{n} w_i(t) x_i(t) + b,$$

kde $i = 0, \ldots, n$, $w_i(t)$ je hodnota váhy ve spojitém čase, b je bias a $y(t)$ je výstupní hodnota ve spojitém čase t. Zde je funkce přechodu F funkcí identity.

Nyní předpokládejme, že vstupní funkce x_i jsou diferencovatelné až do libovolného řádu n.

Uvažujeme lineární diferenciální operátory $L(m, w_n, \dots, w_0)$:

$$\mathbb{C}^n(T) \times \dots \times \mathbb{C}^n(T) \to \mathbb{C}^n(T), \text{ tj. } \mathbb{C}^n(T) \times \dots \times \mathbb{C}^n(T) = [\mathbb{C}^n(T)]^{n+1},$$

definované jako

$$L(m, w_n, \dots, w_0)x(t) = mb + \sum_{k=1}^{n} w_k(t) \frac{d^k x_k(t)}{dt^k},$$

$$x(t) = (x_0(t), x_1(t), \dots, x_n(t)) \in \mathbb{C}^n(T) \times \dots \times \mathbb{C}^n(T) = [\mathbb{C}^n(T)]^{n+1}.$$

Pak označme $\mathbb{L}Ne_n(T)$ jako aditivní abelovskou grupu lineárních diferenciálních operátorů $L(m, w_n, \dots, w_0)$, kde pro $L(m, w_n, \dots, w_0)$, $L(k, w_n^*, \dots, w_0^*) \in \mathbb{L}Ne_n(T)$ s biasem b definujeme

$$L(m, w_n, \dots, w_0) + L(s, w_n^* \dots, w_0^*) = L(m + s, w_n + w_n^*, \dots, w_0 + w_0^*),$$

kde

$$L(m + s, w_n + w_n^*, \dots, w_0 + w_0^*)x(t) = (m + s)b + \sum_{k=0}^{n} (w_k(t) + w_k^*(t)) \frac{d^k x_k(t)}{dt^k},$$

$$t \in T \text{ and } x(t) = (x_0(t), x_1(t), \dots, x_n(t)) \in [\mathbb{C}^n(T)]^{n+1}.$$

Předpokládejme, že $w_k(t) \in \mathbb{C}^n(T)$, a definujme zobrazení

$$\delta_n : \mathbb{C}^n(T) \times \mathbb{L}Ne_n(T) \to \mathbb{C}^n(T)$$

jako

$$\delta_n(x(t), L(m, w_n, \dots, w_0)) = mb + x(t) + m + \sum_{k=0}^{n} w_k(t) \frac{d^k x(t)}{dt^k},\ x(t) \in \mathbb{C}^n(T),$$

kde $w_n, \dots, w_0$ jsou váhy odpovídající vstupům a b je bias neuronu odpovídající operátoru $L(m, w_n, \dots, w_0) \in \mathbb{L}Ne_n(T)$.

Věta 12. *Nechť $\mathbb{L}Ne_n(T)$, $\mathbb{C}^n(T)$ jsou výše definované struktury a $\delta_n : \mathbb{C}^n(T) \times \mathbb{L}Ne_n(T) \to \mathbb{C}^n(T)$ je výše definované zobrazení.Pak trojice $(\mathbb{C}^n(T), \mathbb{L}Ne_n(T), \delta_n)$ je akcí grupy $\mathbb{L}Ne_n(T)$ na grupě $\mathbb{C}^n(T)$, tj. kvazi-automat se stavovým prostorem $\mathbb{C}^n(T)$ a s abecedou $\mathbb{L}Ne_n(T)$ se strukturou grupy umělých neuronů.*

Důkaz. Ověřme podmínku smíšené asociativity (MAC). Předpokládejme, že $x \in \mathbb{C}^n(T)$ a $L(m, w_n, \dots, w_0)$, $L(k, u_n, \dots, u_0) \in \mathbb{L}Ne_n(T)$. Pak máme

$$\delta_n(\delta_n(x(t), L(m, w_n, \dots, w_0), L(k, u_n, \dots, u_0))) =$$

$$= \delta_n(mb + x(t) + m + \sum_{k=0}^{n} w_k(t)\frac{d^k x(t)}{dt^k}, L(k, u_n, \dots, u_0)) =$$

$$= kb + mb + x(t) + m + k + \sum_{k=0}^{n} w_k(t)\frac{d^k x(t)}{dt^k} + \sum_{k=0}^{n} u_k(t)\frac{d^k x(t)}{dt^k} =$$

$$= (m + k)b + x(t) + m + k + \sum_{k=0}^{n}(w_k(t) + u_k(t))\frac{d^k x(t)}{dt^k} =$$

$$= \delta_n(x(t), L(m + k, w_n(t) + u_n(t), \dots, w_0)(t) + u_0)(t)) =$$

$$= \delta_n(x(t), L(m, w_n, \dots, w_0) + L(k, u_n, \dots, u_0)),$$

Takže podmínka MAC je splněna. □

Uvažujme interval $T \subseteq \mathbb{R}$ a okruh $\mathbb{C}(T)$ všech spojitých funkcí definovaných na intervalu. Nechť $\{\varphi_k; k \in \mathbb{N}\}$ je posloupnost endomorfismů okruhu $\mathbb{C}(T)$. Označme $\mathbb{A}_{n+k}\mathbb{N}(T)_m$ jako EL-hypergrupu umělých neuronů konstruovaných výše, s vektory vah dimenze $n + k \in \mathbb{N}(= \{1, 2, 3 \dots\})$. Nechť $[\mathbb{C}(T)]^{n+k} = \mathbb{C}(T) \times \mathbb{C}(T) \times \dots \times \mathbb{C}(T)$ $(n + k$-krát) tj. $\mathbb{C}(T)]^{n+k}$ je $n + k$-dimenzionální kartézská krychle. Označme pomocí $\bar{\varphi}_k : [\mathbb{C}(T)]^{n+k} \to [\mathbb{C}(T)]^{n+k-1}$ extenzi φ_k, takže $\bar{\varphi}_k(\vec{w}) = \bar{\varphi}_k(w_1, \dots, w_{n+k-1}, w_{n+k}) = (w_1, \dots, w_{n+k-1})$. Označme zobrazení $F_k : \mathbb{A}_{n+k}\mathbb{N}(T)_m \to \mathbb{A}_{n+k-1}\mathbb{N}(T)_m$ definované jako $F_k(Ne(\vec{w})) = Ne(\vec{w}_1)$ s $\vec{w}_1 = (w_1, \dots, w_{n+k})$. Vezměme v potaz skryté množiny hypergrup $\mathbb{A}_{n+k}\mathbb{N}(T)_m$ disponující výše definovanou uspořádající relací:

$$\text{pro } \vec{w} = (w_1, \dots, w_{n+k}),\ \vec{u} = (u_1, \dots, u_{n+k}) \in [\mathbb{C}(T)]^{n+k}$$

máme $\vec{w} \le \vec{u}$, když $w_r \le u_r$, $r = 1, 2, \dots, n + k$ a $w_m \le u_m$. Nyní pro $Ne(\vec{w})$, $Ne(\vec{u}) \in \mathbb{A}_{n+k}\mathbb{N}(T)_m$ takové, že $\vec{w} = (w_1, \dots, w_{n+k})$, $\vec{u} = (u_1, \dots, u_{n+k})$, $Ne(\vec{w}) \le Ne(\vec{u})$, což znamená $\vec{w} \le \vec{u}$ ($w_m = u_m$ a biasy korespondujících neuronů jsou stejné) pak máme $\bar{\varphi}_k(\vec{w}) = (w_1, \dots, w_{n+k-1}) \le (u_1, \dots, u_{n+k-1}) = \bar{\varphi}_k(\vec{u})$, což implikuje $F_k(\vec{w}) \le F_k(\vec{u})$.

Následně zobrazení $F_k : (\mathbb{A}_{n+k}\mathbb{N}(T)_m, \le) \to (\mathbb{A}_{n+k-1}\mathbb{N}(T)_m, \le)$ zachovává uspořádání, tj. jedná se o pořádkový homomorfismus hypergrup. Konečným výsledkem našich úvah je následující posloupnost hypergrup umělých neuronů a propojení homomorfismů:

$$\mathbb{AN}_n(T)_m \xleftarrow{F_1} \mathbb{AN}_{n+1}(T)_m \xleftarrow{F_2} \cdots \xleftarrow{F_K} \mathbb{AN}_{n+k}(T)_m \xleftarrow{F_{K+1}} \mathbb{AN}_{n+k+1}(T)_m \dots$$

3.6 Systémy fragmentů umělých neuronových sítí

Z neuronových sítí lze získat některé fragmenty, zejména ve tvaru konečných stromů [69]. Protože stromy a lesy jsou zvláštní uspořádané množiny a v algebraizovaných formách jsou zkoumány podrobně, lze využít některé výsledky týkající se vlastností stromů a lesů a také vlastností monoidů jejich transformací. Některé výsledky se získají pro nekonečné lokálně konečné stromy, tj. stromy obsahující řetězce nekonečných délek, avšak stromy tohoto tvaru mohou být omezeny vhodnými kongruencemi. Transformační stromy jsou založeny na konceptu regularity monoidu jejich lokálních automorfismů, což bylo dalším dílčím cílem výzkumu.

Při konstrukci multistruktur umělých neuronů používáme vhodné zcela přirozené uspořádání, které je definováno na množině $\mathbb{AN}(\mathbf{T})$ odpovídajících umělých neuronů. Mimo to neuronové sítě obsahují i konečné stromy jako své subsystémy.

Poznámka 8. Aby nedošlo k záměně časových řad $\mathbf{T}$ se stejně pojmenovanou stromovou strukturou, je použito v této části práce tučné písmo v názvu časové sady [69]. Více informací o algebraických strukturách umělých neuronů lze nalézt také v [66], [67].

V následující části je konstruované stromové uspořádání v kolekci umělých neuronů $\mathbb{AN}(\mathbf{T})$. Zvažme neurony $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(\mathbf{T})$ definované pomocí funkcí:

$$y_r(t) = \sum_{k=1}^{n} w_k(t)x_k(t) + b_r, \tag{3.5}$$

$$y_s(t) = \sum_{m=1}^{n} w_m(t)x_m(t) + b_s,\ t \in \mathbf{T} \tag{3.6}$$

v daném pořadí, kde symboly r, s označují konečné posloupnosti nezáporných celých čísel. Jestliže $r = [r_1, r_2, \dots, r_k]$, $s = [s_1, s_2, \dots, s_l]$ jsou takové posloupnosti (vektory), pak položíme $r < s$ když $k < l$ a $r_1 = s_1, r_2 = s_2, \dots, r_k = s_k$. Kořen stromu je určen posloupností s jedním prvkem $r = [0]$ nebo $r = [r_v]$, $r_v \in \mathbb{N}$. Pak pro neurony $Ne(\vec{w}_r)$, $Ne(\vec{w}_s)$ podobně definujeme $Ne(\vec{w}_r) < Ne(\vec{w}_s)$ kdykoliv $r < s$.

Charles Wells popisuje v článku [119] mimo jiné lokálně konečné stromy s tranzitivní akcí monoidů jejich lokálních automorfismů (tj. izotonní transformace restrikcí, které jsou na omezených intervalech pořádkovými izomorfismy) a grupy automorfismů (Věta 1) definovaných na těchto stromech. Výrazem netriviální lokálně konečný strom (les ve vhodnější terminologii) se rozumí nejméně dvouprvková částečně uspořádaná množina, jejíž všechny duální hlavní ideály jsou dobře uspořádané mající ordinální číslo maximálně ω tj. ordinální číslo množiny $\mathbb{N}$ všech kladných celých čísel). Při využití výsledků článků [52], [50], [119] a [121] zabývajících se určitými algebraickými a topologickými charakteristikami lokálně konečných a dolních lesů s tranzitivními akcemi lokálních automorfních monoidů na jejich nosných množinách je v dalším využito konstrukce souboru umělých neuronů z předchozích částí práce. V kategorii lokálně konečných stromů a lesů hrají důležitou roli tzv. lokální izomorfismy nebo lokální automorfismy. Tato zobrazení byla popsána Charlesem Wellsem v jeho příspěvku [119] a zkoumána v některých dalších pracích [13, 56, 58]. Důležitou vlastností pologrup a monoidů, která je studována v mnoha dokumentech, je koncept regularity zavedený Johnem von Neumannem pro okruhy v roce 1936 - cf. [21]. To se blíží pojmu pseudoinverzních matic a také retrakční konstrukci struktur. Zejména v článku [14] jsou zkoumány uspořádané regulární pologrupy včetně přirozeně uspořádaných pologrup - s Nambooripadovým uspořádáním - a odkazy zmíněného článku [14] obsahují literaturu věnovanou těmto tématům. Tyto otázky jsou zkoumány v další části textu práce. Význam hlavních vět také spočívá v tom, že ukazují, že za předpokladu tranzitivní akce lokálních automorfních monoidů na lokálně konečných lesích tyto mohou být opatřeny Alexandrovovou topologií [2], [3] (tj. kvazi-diskrétní ve smyslu [25] kap. V.) tak, že lokální automorfní monoid se shoduje s monoidem všech lokálních homeomorfismů nebo s monoidem všech uzavřených spojitých zobrazení do sebe v dotčeném prostoru. Navíc pro jakýkoli pár různých bodů existují topologie s výše zmíněnými vlastnostmi semi-separující tyto body.

Terminologie týkající se stromů používaná v další části práce je známá v literatuře (z různých hledisek) pohledu) - srov. [83], [119]. V souladu s [119] říkáme, že uspořádaná množina $(T, \leq)$ je *horní lokálně konečný les*, pokud všechny jeho duální hlavní ideály jsou dobře uspořádány s ordinálním číslem nanejvýš ω. Říkáme, že $(T, \leq)$ je *dolní lokálně konečný les*, pokud $(T, \leq^{-1})$ je horním lokálně konečným lesem. Souvislý horní (dolní) lokálně konečný les se nazývá *horní (dolní) lokálně konečný strom*. Tím pádem horní lokálně konečný strom je uspořádaná množina, což je horní polosvaz, jehož všechny ohraničené intervaly jsou konečné řetězce. Kořenem horního (dolního) stromu je největší (nejmenší) prvek tohoto stromu. Říkáme, že les je *anti kořenný,* jestliže žádný z jeho maximálních stromů nemá kořen. Pomocí Max $(T, \leq)$ (Min $(T, \leq)$) označujeme množinu všech maximálních (minimálních) prvků $(T, \leq)$. Hlavní (duální) ideál $(T, \leq)$ generovaný prvkem $x \in T$ je označen $(x]_{\leq}$ nebo $(x]$ ($[x)_{\leq}$ nebo $[x)$). Interval s počátečním prvkem s a koncovým prvkem t, tj. množina $\{x | s \leq x \leq t\}$ je označena pomocí $[s, t]$; $s \prec t$, je míněno $[s, t] = \{s, t\}$.

Izotonní zobrazení f lesa (stromu) $(T, \leq)$ do sebe je považováno za *lokální automorfismus* $(T, \leq)$, pokud pro jakýkoli pár prvků $s, t \in T$, $s < t$, restrikce $f|[s, t]$ je izomorfismus řetězce $[s, t]$ na řetězec $[f(s), f(t)]$. Monoid všech místní automorfismů lesa $(T, \leq)$ bude označen $\mathrm{LA}(T, \leq)$. Jako obvykle o monoidu S (s jednotkou e) se říká, že má akci na množině X, pokud je dáno zobrazení $\pi : S \times X \to X$, takže $\pi(e, x) = x$, $\pi(s_1 s_2, x) = \pi(s_1, \pi(s_2, x))$ pro všechna $x \in X$, s_1, $s_2 \in S$. Pro submonoid F symetrického monoidu $(X^X, \cdot)$ z X položíme $\pi(f, x) = f(x)$ pro $f \in F$, $x \in X$. Monoid F *operuje tranzitivně* na X, pokud pro jakýkoli pár prvků $x_1, x_2 \in X$, existuje zobrazení $f \in F$, takové, že $f(x_1) = x_2$.

Nechť f je transformace - tj. zobrazení do sebe množiny T. Monoid $C_T(f) = \{g \in T^T | fg = gf\}$ se nazývá centralizátor transformace f (v symetrický monoid $(T^T, \cdot)$). Ve shodě s [6] lze definovat binární operaci $\circ$ na lokálně konečném stromu

$(T, \leq)$ následovně.

Předpokládejme, že $(T, \leq)$ je horní lokálně konečný strom. Pro $s, t \in T$ položíme $\delta(s,t) = \mathrm{card}[t, \sup\{s,t\}] - \mathrm{card}[s, \sup\{s,t\}]$ (oba kardinály na pravé straně jsou konečné). Pak položíme $s \circ t = t^+$ (kde t^+ je následníkem t), pokud $\delta(s,t) \geq 0$ a $t \notin \mathrm{Max}(T, \leq)$. Pokud je t největším prvkem $(T, \leq)$, pak $s = t$ a položíme $s \circ t = t$. Dále $s \circ t = s^+$ kdykoli $\delta(s,t) < 0$. Pokud $(T, \leq)$ je dolní lokálně konečný strom, bereme v úvahu strom $(T, \leq^{-1})$ a násobení jeho prvků, jež je definováno s ohledem na uspořádání $\leq^{-1}$, které vytváří strukturu horního stromu na T. Monoid endomorfismů grupoidu $(T_i, \circ)$ je označen jako $\mathrm{End}(T_i, \circ)$.

Pojem (konečné) stromové algebry představil Ladislav Nebeský v [89]. Zobecnění tohoto pojmu pro případ nekonečných nosičů zkoumá Bohdan Zelinka v [121], kde mimo jiné charakterizuje stromové algebry realizované (neorientovanými) stromovými grafy. Nejprve popišme některé potřebné termíny z [121]. *Stromová algebra* $\mathcal{A} = (T, P)$ je algebra s nosičem (nosnou množinou) T a s jednou ternární operací P, která splňuje následující podmínky pro libovolné prvky t, u, v, w z T :

1. $P(t, t, u) = t$,
2. $P(t, u, v) = P(u, t, v) = P(t, v, u)$,
3. $P(P(t, u, v), u, w) = P(t, u, P(v, u, w))$,
4. $P(t, u, w) \neq P(u, v, w) \neq P(t, v, w) \Rightarrow P(t, u, w) = P(t, v, w)$.

Nechť $\mathcal{A} = (T, P)$ je stromová algebra, $t, u \in T$. Omezený interval $\mathcal{A}$ určený t a u je množina $S(t, u) = \{x \in T | P(t, u, x) = x\}$. Nechť $x_0, x_1, x_2, \ldots$ je nekonečná posloupnost po dvou různých prvků T s těmito vlastnostmi: $S(x_0, x_i) \subset S(x_0, x_{i+1})$ pro každé kladné celé číslo i a pro každé $y \in T$ existuje kladné celé číslo i tak, že $x_i \notin S(x_0, y)$. Pak množina $D_{x_0} = \bigcup_{i=1}^{\infty} S(x_0, x_i)$ se nazývá neohraničený interval $\mathcal{A}$ s počátečním prvkem x_0. Stromová algebra $\mathcal{A}$ se nazývá diskrétní, pokud interval $S(u, v)$ v $\mathcal{A}$ pro jakékoli dva prvky t, u jeho nosiče je konečný. Věta 12 z [121] říká, že existuje vzájemná korespondence mezi přesně diskrétními stromovými algebrami a stromovými grafy $G = (T, H)$ daná pravidlem: $P(t, u, v)$ je (jediný) společný vrchol cesty v G spojující t a u, cesty v G spojující t a v a cesty v G spojující u a v.

Nechť $(T, \leq)$ je lokálně konečný les ve smyslu výše uvedené definice. Označme pomocí $(T, \rho_{\leq})$ reflexi orientovaného grafu (v kategorii orientovaných grafů a homomorfismů) určené symetrizací, tj. $t\rho_{\leq}s$ právě tehdy, když $t \leq s$ nebo $s \leq t$. Jestliže $\{(T_i, \leq) | i \in I\}$ je soubor všech maximálních stromů lesa $(T, \leq)$, pak pomocí $\mathcal{A}_i = (T_i, P_i)$, pro $i \in I$, označujeme diskrétní algebru stromů odpovídající (neorientovanému) stromovému grafu $(T_i, \rho_{\leq})$. Předpokládejme, že $\mathcal{A} = (T, P)$ obsahuje alespoň jeden neohraničený interval D_t a označme $E(D_t)$ konec $\mathcal{A}$ určený D_t, tj. množinu všech neohraničených intervalů $\mathcal{A}$ tak, že průsečík z každého z nich s D_t je také neohraničený interval. Definujme na $E(D_t)$ ternární operaci P^* tímto způsobem: $P^*(D_u, D_v, D_w) = D_u \cap D_v \cap D_w$ pro jakoukoli trojici D_u, D_v, D_w prvků z $E(D_t)$. Je snadné ověřit, že $(E(D_t), P^*)$ je stromová algebra nazývaná koncová stromová algebra určená neohraničeným intervalem D_t algebry $\mathcal{A}$ - izomorfní s algebrou $\mathcal{A}$; odpovídající izomorfismus $\varphi : (T, P) \to (E(D_t), P^*)$ je definován jako $\varphi(x) = D_x \in E(D_t)$ pro libovolné $x \in T$.

Bližší informace o teorii retraktů pro uspořádané množiny lze nalézt v [105] a dalších referencích citovaných tamtéž. Podmnožina $\mathcal{Q}$ uspořádané množiny $(S, \leq)$ je retrakce $(S, \leq)$, pokud existuje zobrazení zachovávající uspořádání $g : (S, \leq) \to (\mathcal{Q}, \leq)$ (nazývá se retrakce), což je identické zobrazení na $\mathcal{Q}$. V souladu s tím představa o uspořádané podmnožině $(A, \leq)$ lokálně konečného lesa $(T, \leq)$ je považována za LA-*retrakci* lesa $(T, \leq)$, pokud existuje lokální automorfismus $g : (T, \leq) \to (A, \leq)$ – nazývá se LA-*retrakcí* – takovou, že $g|A = id_A$. Je třeba poznamenat, že některé výsledky z teorie retraktů uspořádaných množin lze přenést na uvažovaný případ, např. každý maximální řetězec lokálně konečného stromu je jeho LA-retrakcí (srov. [105], str. 104). Minimalita LA-retrakce znamená minimální vzhledem k přirozenému uspořádání podle množinové inkluze.

Alexandrovovým topologickým prostorem (nazývaným také kvazi-diskrétní) máme na mysli ve shodě s [2], [3] dvojici (X, τ), kde X je množina a τ je zcela aditivní topologická uzávěrová operace na X. Kvazi-diskrétní T_0-prostory byly podrobněji prozkoumány pravděpodobně poprvé v [5]; odtud vyplývá užitečnost pojmu spojitých uzavřených zobrazení (zjednodušené zobrazení) takových prostorů. Monoid všech uzavřených deformací (tj. uzavřených spojitých zobrazení do sebe) topologického prostoru (X, τ) s obvyklou kompozicí zobrazení jako násobení je označeno $S(X, \tau)$. Zobrazení f topologického prostoru (X, τ) do (Y, σ) se nazývá lokální homeomorfismus, pokud pro jakýkoli bod $x \in X$ existuje τ-okolí O_x bodu x tak, že restrikce $f|O_x$ je homeomorfismus O_x na $f(O_x)$. Monoid všech lokálních homeomorfismů (X, τ) do sebe značíme $\mathrm{LH}(X, \tau)$. Body x, y topologického prostoru se nazývají semiseparované ([25]) pokud existují okolí O_x, O_y bodů x, y, takové, že $y \notin O_x$ a $x \notin O_y$. Pokud navíc $O_x \cap O_y = \emptyset$, pak body x, y se nazývají oddělené. Topologický prostor je považován za perfektní, pokud neobsahuje izolované body. Lokálně konečné stromy a lesy lze získat ze struktur neuronových sítí - viz níže.

Soubor všech kladných celých čísel je označen jako $\mathbb{N}$ a $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Základní pojmy teorie pologrup lze nalézt v [21], pojmy týkající se částečně uspořádaných množin (nazývá se krátce v anglických zdrojích „posets“) v [12]. Některé z nich jsou důležité pro následující konstrukce. Takže (vzhledem k [12]) je délka konečného řetězce $K = \{x_1, \ldots, x_n\}$ typu $n \in \mathbb{N}$ definována jako $l(K) = n - 1$. Obecněji je délka $l(P)$ částečně uspořádané množiny P definována jako supremum (nejmenší horní hranice) délek řetězců v P. Lokálně konečnou částečně uspořádanou množinou máme na mysli tříděnou částečně uspořádanou množinu ve smyslu [12, s. 5], tj. jakýkoli řetězec mezi libovolnými dvěma srovnatelnými prvky je konečný. Viz také zdroj [111] věnovaný určitým transformačním pologrupám částečně uspořádaných množin, které splňují podmínku, že mezi každou dvojicí srovnatelných prvků existuje nasycený řetězec spojující tyto prvky. Izotonní zobrazení počátku do sebe se nazývá endomorfismus ve shodě s [4]. Lokální automorfismus částečně uspořádané množiny $(M, \leq)$ je endomorfismus f množiny $(M, \leq)$ takový, že pro jakýkoli pár $a, b \in M$, $a \leq b$ restrikce $f|[a, b]_{\leq}$ je pořádkový izomorfismus $[a, b]_{\leq} (= \{x|a \leq x \leq b\})$ na interval $[f(a), f(b)]_{\leq}$ (cf. [119] pro lokálně konečné stromy a lesy). Monoid všech endomorfismů (lokálních automorfismů) částečně uspořádané množiny $(M, \leq)$ je označen jako $\mathrm{End}(M, \leq)$ $(\mathrm{LA}(M, \leq))$. V souladu s teorií retraktů- (srov. [105]) definujeme LA-retrakci částečně uspořádané množiny $(M, \leq)$ jako podmnožinu A množiny M tak, že zde existuje lokální automorfismus $g : (M, \leq) \to (A, \leq)$ – nazývaný LA-retrakce – s vlastností $g|A = id_A$.

Pro $S \subset M$ označujeme pomocí $(S]_{\leq}$ $([S)_{\leq})$ ideál (duální ideál) množiny $(M, \leq)$ generovaný S. Ideály (duální ideály) částečně uspořádaných množin se v literatuře nazývají také dolní množiny (horní množiny). Pokud je S singleton, pak se odpovídající ideály nazývají hlavní ideál a píšeme např. $(x]_{\leq}$ namísto $(\{x\}]_{\leq}$. Lokálně konečným stromem máme na mysli souvislou lokálně konečnou částečně uspořádanou množinu, jejíž každý hlavní ideál je buď konečný řetězec (kořenový lokálně konečný strom), nebo řetězec typu ω^d (zde $\leq^d$ znamená duální uspořádání k $\leq$). Jak již bylo výše popsáno, lokálně konečný kořenový strom je také úplný dolní polosvaz s ohledem na dané uspořádání. Další informace o lokálních automorfních monoidech lokálně konečných stromů a lesů jsou zpracovány v [119], [13] a [58]. Jestliže prvek b pokrývá prvek a, píšeme $[a, b]_{\leq} = \{a, b\}$ (tyto dvouprvkové řetězce se také nazývají primární řetězce). Pokud je x prvkem lokálně konečného stromu, pak x^- je jedinečný prvek pokrytý x (předchůdcem x), zatímco x^+ je množina všech následovníků x, tj. $x^+ = \{y|x \prec y\}$. Množina všech maximálních (minimálních) prvků $(M, \leq)$ je označena $\mathrm{Max}(M, \leq)$, $\mathrm{Min}(M, \leq)$. Symbol $\oplus$ se používá pro ordinální sumu.

Plně transformační (nebo symetrický) monoid množiny M je označen jako obvykle T_M. Používáme klasický výsledek C. G Dosse ([21], §1.9, cvičení 1) který říká, že T_M je vždy regulární.

Charakterizace neprázdných částečně uspořádaných množin (které nejsou řetězci) běžnými endomorfními monoidy je podle A. J. Aizenštatové [4]. Pokud jde o řetězce, viz §1 uvedeného zdroje.

Věta 13. *([4], Věta 2.6) Buď $(M, \leq)$ neprázdná částečně uspořádaná množina, která není ani řetězcem, ani protiřetězcem. Pak* $\mathrm{End}(M, \leq)$ *je regulární, pokud $(M, \leq)$ splňuje alespoň jednu z následujících podmínek:*

1. *Existuje pár protiřetězců M_1, M_2, takže $(M, \leq) = M_1 \oplus M_2$.*

2. *Existuje dvouprvkový rozklad M_1, M_2 množiny M a pár prvků $(a,b) \in M_1 \times M_2$ tak, že pro $x, y \in M$ máme $x < y$ právě tehdy, když buď $x = a$, $y \in M_2$, nebo $x \in M_1$ a $y = b$.*
3. *$M = \{a_1, a_2, a_3, b_1, b_2, b_3\}$, $a_1 < b_1$, $a_2 < b_1$, $a_1 < b_2$, $a_2 < b_3$, $a_3 < b_2$, $a_3 < b_3$ (tj. $(M, \leq)$ je koruna se šesti prvky).*
4. *Existuje pár prvků $(a,b) \in M$, takže $M = \{a\} \oplus M_0 \oplus \{b\}$, kde $M_0 = M \setminus \{a,b\} \neq \emptyset$ je protiřetězec.*

Okamžitým důsledkem věty 13 je následující tvrzení:

Tvrzení 8. *Nechť $(T, \leq)$ je strom, který není řetězcem, takže* $\mathrm{End}(T, \leq)$ *je regulární. Pak $(T, \leq)$ je lokálně konečný kořenový strom délky 1.*

V právě zmíněném případě máme $\mathrm{LA}(T, \leq)$, který je isomorfní k symetrické grupě množiny $\mathrm{Max}(T, \leq)$. Nyní charakterizujeme lokálně konečný strom, jehož lokálního automorfní monoid je regulární. S použitím výsledku L. A. Skornjakova ([109], Věta 1) dostaneme (viz také [13] Věta 2.l):

Tvrzení 9. *Nechť $(T, \leq)$ je lokálně konečný strom bez kořene. Pak* $\mathrm{LA}(T, \leq)$ *je regulární právě tehdy, když $T = K \cup \mathrm{Max}(T, \leq)$, kde $(K, \leq)$ je řetězec typu $\omega^d \oplus \omega$.*

Poznámka 9. V případě kořenových lokálně konečných stromů Věta 1 z [109] není použitelná pro rozdíl mezi lokálními automorfismy kořenového stromu a endomorfismy monounární algebry daných funkcí předchůdce. Použijeme metodu dvou invariantů - jmenovitě stupně a úrovně funkcí. Funkce stupně byla zavedena M. Novotným v [98], [99] za účelem konstrukce homomorfismů monounárních algeber. Pro některé aplikace viz např. [100] a [101].

V dalším textu $(T, \leq)$ znamená kořenový lokálně konečný strom s kořenem (nejmenší prvek) t_*. Bude se dále nazývat stromem. Označme Ord třídu všech ordinálů, ∞ symbol, který nepatří do Ord. Položíme $\alpha < \infty$ pro všechny ordinály α. Dále T^∞ je podmnožinou všech $t \in T$, takže existuje řetězec $K \subset [t)_\leq$ typu ω, $T^0 = \mathrm{Max}(T, \leq)$. Nechť je $\alpha \in Ord$, $\alpha > 0$ a předpokládejme, že T^β je definováno pro všechny $\beta < \alpha$. Pak položíme

$$T^\alpha = \{t \in T \setminus (T^\infty \cup \bigcup_{\alpha<\beta} T^\beta) | t^+ \subset \bigcup_{\alpha<\beta} T^\beta\}. \tag{3.7}$$

Stejně jako v [98], [99] lze prokázat, že existuje nejmenší ordinál ϑ_T, takže $T^\varkappa = \emptyset$ pro všechny $\varkappa \geq \vartheta_T$ a $T = T^\infty \cup \bigcup_{\alpha<\vartheta_T} T^\alpha$ s nesouvislými sčítanci ([98], Lemma 2.3 nebo [99], §3). Pak definujeme ordinální funkci $d_T : T \to Ord \cup \{\infty\}$ jako $d_T(t) = \alpha$ pro všechny $t \in T^\alpha$. Ordinál $d_T(t)$ je označován jako stupeň prvku t v $(T, \leq)$. Připomeňme si, že výšková funkce $h_T : T \to \mathbb{N}_0$ je definována takto: $h_T(t_*) = 0$

a $h_T(s) = h_T(t) + 1$ pro každé $s \in t^+$ a libovolné $t \in T$. Nezáporné celé číslo $h_T(t)$ je považováno za výšku $t \in T$.

Z definice funkce d_T snadno vyplývá, že $d_T(T) \subset \mathbb{N}_0$ implikuje existenci celého čísla $k \in \mathbb{N}_0$, takže $d_T(t) \leq k$ pro libovolné $t \in T$, tj. $d_T(t_*) \in \mathbb{N}_0$. To odpovídá skutečnosti, že strom $(T, \leq)$ má konečnou výšku (tj. má konečnou délku jako uspořádaná množina - [12], I., 3). Proto je výška lokálně konečného kořenového stromu $(T, \leq)$ buď $\mathrm{Max}\{h_T(t) | t \in T\}$, pokud existuje maximum, nebo ω v opačném případě. Pro libovolné $t \in T$ položíme

$$L(t) = \{s \in T | d_T(s) = d_T(t)\}, \tag{3.8}$$

$L(t)$ se nazývá úroveň určená prvkem t. Je zřejmé, že existuje nejvýše spočetná podmnožina $S \subset T$ tak, že

$$T = \bigcup_{t \in S} L(t) \tag{3.9}$$

s nesouvislými sčítanci a v případě konečného stromu $(T, \leq)$, tento rozklad se shoduje s úrovní rozkladu stromu uvažovaného v článku [32].

Pomocí funkcí d_T, h_T definujeme binární relaci σ_T a tři ternární vztahy ϱ_T, $\varrho_T^{'}$, $\varrho_T^{''}$ na množině T takto:

Dvojice $(t_1, t_2) \in T \times T$ patří k σ_T právě tehdy, když $t_1 \neq t_2$, $h_T(t_1) = h_T(t_2)$ a $d_T(s) = d_T(t)$ pro libovolný pár $(s, t) \in (t_1]_{\leq} \times (t_1]_{\leq}$ uzlů stejné výšky (tj.: $h_T(s) = h_T(t)$).

Dále řekneme, že trojice $(t_1, t_2, t_3) \in T \times T \times T$ splňuje podmínku:

A pokud $t_i \neq t_j$ pro $i \neq j$ a $h_T(t_i) = h_T(t_j)$ pro libovolný pár indexů $i, j \in \{1, 2, 3\}$,

B pokud pro libovolnou trojici $(s_1, s_2, s_3) \in \prod_{i=1}^{3} (t_i]_{\leq}$ uzlů stejné výšky máme

$$d_T(s_1) \leq d_T(s_2) \leq d_T(s_3),$$

B$^{'}$ jestliže $(s, t) \in \sigma_T$ pro nějaký pár $(s, t) \in (t_1]_{\leq} \times (t_1]_{\leq}$ uzlů stejné výšky,

B$^{''}$ pokud pro $u \in (t_1]_{\leq} \cap L(t_2 \wedge t_3)$ zde existují uzly $a, b \in L(t_1) \cap (u]_{\leq}$, takže $t_1 \wedge a \in (u, t_1]_{\leq}$, $(a, t_2) \in \sigma_T$, $d_T(t_2) < d_T(b)$ a $\{t \in [t_2 \wedge t_3)_{\leq}, (b, t) \in \sigma_T\} = \{t_3\}$,

C pokud $t_1 \wedge t_2 = t_1 \wedge t_3$,

C$^{'}$ jestliže $t_1 \wedge t_2 = t_2$

C$^{''}$, pokud $t_1 \wedge t_2 < t_2 \wedge t_3$.

Nyní definujeme ternární relace ϱ_T, $\varrho_T^{'}$, $\varrho_T^{''}$ tímto způsobem:

Trojice (t_1, t_2, t_3) patří do ϱ_T právě tehdy, když splňuje (A), (B), (C), patří do $\varrho_T^{'}$ právě tehdy, když splňuje (A), (B$^{'}$), (C$^{'}$), a nakonec $(t_1, t_2, t_3) \in \varrho_T^{''}$ právě tehdy, když tato trojice vyhovuje (A), (B$^{''}$), (C$^{''}$).

Charakterizační věta o regularitě $\mathrm{LA}(T, \leq)$ bude formulována pomocí relací ϱ_T, $\varrho_T^{'}$, $\varrho_T^{''}$.

Poznámka 10. Z definice funkcí d_T, h_T vyplývá, že pro pár $s, t \in T$ existuje $f \in \mathrm{LA}(T, \leq)$ taková, že $f(s) = t$ právě tehdy, když $h_T(s) = h_T(t)$ a $d_T(u) \leq d_T(v)$ pro libovolný pár $(u, v) \in (s]_{\leq} \times (t]_{\leq}$ prvků stejné výšky (srov. [100], str. 172, podmínka (C^1) nebo věta v [90]).

Lemma 3. *Nechť $(T, \leq)$ je strom $(t_1, t_2, t_3) \in T \times T \times T$ je trojice vzájemně odlišných uzlů stejné výšky. Následující tvrzení jsou ekvivalentní:*

1. *Existuje $f \in \mathrm{LA}(T, \leq)$, takže $f(t_i) = t_{i+1}$ pro $i = 1, 2$.*
2. *Buď $(t_1, t_2, t_3) \in \varrho_T$ nebo $(s, t) \in \sigma_T$ pro pár $(s, t) \in (t_2]_{\leq} \times (t_3]_{\leq}$ a podmínky (B), ($\mathrm{C}^{'}$) jsou splněny .*

Důkaz tohoto Lemma 3 najdete v [56].

Lemma 4. *Pokud je monoid $\mathrm{LA}(T, \leq)$ regulární a trojice $(t_1, t_2, t_3) \in \varrho_T$ splňuje podmínku ($\mathrm{C}^{'}$), pak splňuje také podmínku ($\mathrm{B}^{'}$).*

Důkaz. Pokud $(t_1, t_2, t_3) \in \varrho_T$, pak podle Lemma 4 existuje $f \in \mathrm{LA}(T, \leq)$ tak, že $f(t_i) = t_{i+1}$, $i = 1, 2$. Z odpovídající modifikace konstrukce K (Poznámka 10) zde vyplývá že f lze definovat tak, že

$$f^{-1}(t_2) \subset \mathrm{L}(t_2) \cap (t_1 \wedge t_2, t_1]_{\leq}. \tag{3.10}$$

Protože $f \in \mathrm{LA}(T, \leq)$ je regulární, máme $fgf = f$ pro některé $g \in \mathrm{LA}(T, \leq)$. Označme $s^{'} = t_1 \wedge g(t_2)$. Potom podle 3.10 $s^{'} \in (t_1 \wedge t_2, t_1]_{\leq}$. Nechť $t^{'}$ je prvkem $\mathrm{L}(s^{'}) \cap (t_2]_{\leq}$ (průsečík je singleton). Protože $f([t_1 \wedge t_2, g(t_2)]_{\leq})$ a $g([t_1 \wedge t_2, t_2]_{\leq}) = [t_1 \wedge t_2, g(t_2)]_{\leq}$ kde $f(s^{'}) = t^{'}$, $g(t^{'}) = s^{'}$ máme $d_T(s) = d_T(t)$ pro libovolný pár $(s, t) \in (g(t_2)]_{\leq} \times (t_2]_{\leq}$ uzlů stejné výšky, tedy $(s^{'}, t^{'}) \in \sigma_T$, tj. trojice (t_1, t_2, t_3) splňuje podmínku ($\mathrm{B}^{'}$). □

Lemma 5. *Nechť $(T, \leq)$ je strom konečné výšky, takže $\varrho_T \subset \varrho_T^{'} \cup \varrho_T^{''}$. Potom pro každý uzel $t \in T$ platí následující tvrzení:*

1. *Restrikce $d_T | t^+$ je nejvýše dvouhodnotová.*
2. *Pokud $d_T | t^+$ je dvouhodnotová, pak $d_T(s_0) = \max\{d_T(s) | s \in t^+\}$ pro přesně jeden uzel $s_0 \in t^+$.*

Důkaz.

1. Nechť $t \in T$ je libovolný uzel, $t_i \in t^+$, $i = 1, 2, 3$, je trojicí vzájemně odlišných uzlů, takže $d_T(t_i) \leq d_T(t_{i+1})$ pro $i = 1, 2$. Pak $(t_1, t_2, t_3) \in \varrho_T \setminus \varrho_T^{''}$, (t_1, t_2, t_3) vyhovuje ($\mathrm{B}^{'}$), což znamená $d_T(t_1) = d_T(t_2)$.
2. Předpokládejme $t_i \in t^+$, $i = 1, 2, 3$ opět, $d_T(t_1) < d_T(t_j)$, $j = 2, 3$. Stejným způsobem jako výše $t_2 \neq t_3$ implikuje $d_T(t_1) = d_T(t_2)$, tedy $t_2 = t_3$.

□

Hlavním výsledkem uvedeným v tomto odstavci je následující věta:

Věta 14. *Nechť* $(T, \leq)$ *je lokálně konečný kořenový strom. Monoid* $\mathrm{LA}(T, \leq)$ *je regulární právě tehdy, když má strom konečnou výšku a relace* $\varrho_T^{'} \cup \varrho_T^{''}$ *je extenze relace* ϱ_T*, tj.* $\varrho_T \subset \varrho_T^{'} \cup \varrho_T^{''}$.

Důkaz této věty 14 lze nalézt v [56].

Pro ilustraci uvedených zjištění konkrétními příklady lze použít mimo jiné stromové uspořádání jazyků nad konečnými abecedami. Takže vzhledem ke konečné množině A nazvané abeceda označujeme A^* (jako obvykle) volný monoid slov nad A (s prázdným slovem Λ jako jednotkou). Zvažme následující uspořádání $\leq$ na A^*: Pro $u, v \in A^*$ položíme $u \leq v$ právě tehdy, když $uw = v$ pro nějaké $w \in A^*$. Je zřejmé, že jakýkoli jazyk $T \subset A^*$, $T \neq \emptyset$, který je ideálem částečně uspořádané množiny $(A^*, \leq)$ je kořenový lokálně konečný strom s kořenem Λ. Poznamenejme, že místo $aa \ldots a$ (n- krát) píšeme a^n.

Příklady.

1. Je zřejmé, že binární strom $(\{0,1\}^*, \leq)$ všech konečných slov nad dvouprvkovou abecedou $\{0,1\}$ nemá regulární monoid všech lokálních automorfismů. Monoid $\mathrm{LA}(T^n, \leq)$, kde $T^n = \bigcup\limits_{0 \leq k < n} L_k$, $L_0 = \{\Lambda\}$ a L_k je množina všech slov o délce k tvořená 0 a 1, je regulární tehdy a jen tehdy, když $n \leq 1$.
2. a Nechť $n \geq 2$ je celé číslo, $\{(K_i, \leq); 0 \leq i < \omega\}$ je soubor konečných disjunktních řetězců, takže K_0 má typ $n+1$ a K_j, $j > 1$, má typ n. Předpokládejme $0 \notin \bigcup\limits_{0 \leq i < n} K_i$ a položme $(T_1, \leq) = \{0\} \oplus [(K_0, \leq) + \sum\limits_{1 \leq i < \omega} (K_i, \leq)]$, kde symboly $+, \sum$ znamená kardinální částku (přes antiřetězce).

   b Pro jakékoli celé číslo i, takže $1 \leq i < \omega$ předpokládáme $i \| x$, $K_i \leq x$ pro libovolné $x \in K_i$ a navíc $i \notin K_j$ pro libovolné j s vlastností $1 \leq j < \omega$. Označujeme $(T_2, \leq) = \{0\} \oplus \sum\limits_{1 \leq i < \omega} (\{K_i\} \oplus (\{i\} + K_i))$.

   c Položme $T_3 = [0, n] \subset R$ a definujme funkci $f : T \to T$ tímto způsobem: $f(x) = [x]$ (celá část čísla x) pro všechna $x \in T \setminus \mathbb{N}$ a $f(k) = k - 1$ pro libovolné celé číslo $k, 0 < k \leq n$. Dále pro $r, s \in T$ položíme $r \leq s$, pokud $f^m(s) = r$ pro nějaké nezáporné celé číslo m. Pak $(T_k, \leq)$, $k = 1, 2, 3$ jsou diskrétní kořenové stromy výšky n, takže $\mathrm{card} T_1 = \mathrm{card} T_2 = \aleph_0$, $\mathrm{card} T_3 = 2^{\aleph_0}$, které mají regulární monoidy lokálních automorfismů.

Poznámka 11. Použitím věty 2.4. z [81] získáme okamžitě následující charakterizaci regulárních prvků $\mathrm{LA}(T, \leq)$, což lze také přímo dokázat: Transformace $f \in \mathrm{LA}(T, \leq)$ je regulární prvek monoidu $\mathrm{LA}(T, \leq)$ právě tehdy, když $(i) f(T)$ je LA-retrakt $(T, \leq)$, (ii) existuje podstrom $(T_1, \leq)$ stromu $(T, \leq)$ tak, že restrikce $f|T_1 : (T_1, \leq) \to (f(T), \leq)$ je isomorfismus.

3.6.1 Charakterizace tranzitivní akce

Hlavní výsledky předchozích úvah popsaných v úvodu kapitoly jsou formulovány v následujícím textu [69].

Věta 15. *Nechť $(T, \leq)$ je lokálně konečný horní nebo dolní les, $\{(T_i, \leq)|i \in I\}$ buď soubor všech maximálních stromů $(T, \leq)$. Následující podmínky jsou ekvivalentní:*

1. *Monoid $\mathrm{LA}(T, \leq)$ operuje tranzitivně na množině T.*
2. *Uspořádaná množina $(T, \leq)$ nemá žádné maximální nebo minimální prvky.*
3. *Pro každé $i \in I$ a každý prvek t z diskrétní stromové algebry $\mathcal{A}_i = (T_i, P_i)$ existují $\leq$ - srovnatelné prvky $x, y \in T_i$, $x \neq t \neq y$, takže $P_i(x, y, t) = t$.*
4. *Pro libovolný $i \in I$ a jakýkoli prvek $t \in T_i$ existují alespoň dva různé koncové stromové algebry $(E(D_t), P^*)$, $(E(D_t^{'}), P^*)$ tak, že $D_t^{'} \cup D_t$ je minimální LA-retrakt stromu $(T_i, \leq)$*
5. *Pro jakékoli $i \in I$, $(T_i, \circ)$ je jednoduchý[7] grupoid s vlastností $\mathrm{End}(T_i, \circ) = \mathrm{LA}(T_i, \leq)$.*

Věta 16. *Nechť $\mathcal{A} = (M, P)$ je diskrétní stromová algebra, $\leq$ je uspořádání na množině M takové, aby $(M, \leq)$ byla horní nebo dolní polosvaz. Grupa automorfismů $\mathrm{Aut}(M, \leq)$ operuje tranzitivně na M, pokud jsou splněny následující dvě podmínky:*

(i) *Pro libovolný prvek $t \in M$ existují $\leq$ - porovnatelné prvky $x, y \in M$, $x \neq t \neq y$ takové, že $P(x, y, t) = t$.*

(ii) *Pro každou dvojici prvků $t, u \in M$ máme $\mathrm{card}\{x|x \prec t\} = \mathrm{card}\{x|x \prec u\}$ pokud $(M, \leq)$ je horní polosvaz, nebo $\mathrm{card}\{x|t \prec x\} = \mathrm{card}\{x|u \prec x\}$ pokud $(M, \leq)$ je dolní polosvaz.*

Věta 17. *Nechť $(T, \leq)$ je lokálně konečný horní nebo dolní antikořenný les, $\{(Ti, \leq)|i \in I\}$ je soubor všech maximálních stromů lesa $(T, \leq)$. Následující podmínky jsou ekvivalentní:*

1. *Monoid $\mathrm{LA}(T, \leq)$ operuje tranzitivně na množině T.*
2. *Na množině T existuje perfektní Alexandrovova topologie τ, taková, že $\mathrm{LA}(T, \leq) = \mathrm{S}(T, \tau)$.*
3. *Pro každou dvojici různých prvků $a, b \in T$ existuje perfektní Alexandrovova topologie $\tau_{a,b}$ na množině T, takže $\mathrm{LA}(T, \leq) = \mathrm{S}(T, \tau_{a,b})$ a body a, b jsou semiseparované v prostoru $(T, \tau_{a,b})$.*
4. *Existuje Alexandrovova topologie na množině T, takže $\mathrm{LA}(T, \leq) = \mathrm{LH}(T, \sigma)$ a kardinalita σ -uzávěru jakékoli neprázdné podmnožiny T je nekonečná.*
5. *Pro jakýkoli pár různých prvků $a, b \in T$ existuje Alexandrovova topologie $\sigma_{a,b}$ na T s vlastnostmi z podmínek 4 a takové, že body a, b jsou semiseparované v prostoru $(T, \sigma_{a,b})$.*

[7] Grupoid, který nemá vlastní ideály [88].

Poznámka 12. Monoid lokálních homeomorfismů $\mathrm{LH}(T,\sigma)$ z podmínek 4, 5 věty 17 může být nahrazen monoidem otevřených deformací (tj. otevřených spojitých zobrazení uvažovaného prostoru do sebe) a tvrzení zůstávají platná. Nyní ukážeme, že věta 17 je oprávněná, tj. použité pojmy – jako uzavřená deformace, otevřená deformace, místní homeomorfismus se neshodují v případě Alexandrovových prostorů – tj. kvazi-diskrétních prostorů. Zvažme množinu všech nezáporných celých čísel $\mathbb{N}$ a topologii levého uspořádání τ^- na $\mathbb{N}$; takže uzávěry singletonů $\{n\} \subset \mathbb{N}$ jsou množiny $\tau^-\{n\} = \{n, n+1, n+2, \dots\}$. Definujme zobrazení $f : \mathbb{N} \to \mathbb{N}$ položením $f(0) = 1, f(n) = n$ pro $n > 1$. Zřejmě $f \in S(\mathbb{N}, \tau^-), f \notin \mathrm{LH}(\mathbb{N}, \tau^-)$ a f není otevřená deformace prostoru $(\mathbb{N}, \tau^-)$ také. Na druhou stranu f je otevřená deformace duálního prostoru $(\mathbb{N}, \tau^+)$, $f \notin S(\mathbb{N}, \tau^+)$ a současně $f \notin \mathrm{LH}(\mathbb{N}, \tau^+)$. Dále položme $g(n) = n+1$ pro každé $n \in \mathbb{N}$. Zobrazení g je topologické vložení prostoru $(\mathbb{N}, \tau^-)$ do sebe, proto $g \in \mathrm{LH}(\mathbb{N}, \tau^-) \cap \mathrm{LH}(\mathbb{N}, \tau^+)$. Avšak g není otevřenou deformací prostoru $(\mathbb{N}, \tau^-)$ a $g \notin S(\mathbb{N}, \tau^+)$.

3.6.2 Důkazy charakterizačních vět

Důkazy vět uvedené v předchozím odstavci jsou rozděleny do posloupnost důkazů pomocných lemmat.

Lemma 6. *Nechť $(T, \leq)$ je horní nebo dolní lokálně konečný les. Monoid $\mathrm{LA}(T, \leq)$ operuje na množině T tranzitivně právě tehdy, když $\mathrm{Max}(T, \leq) \cup \mathrm{Min}(T, \leq) = \emptyset$.*

Důkaz. Toto tvrzení bezprostředně vyplývá z [25] -tvrzení (a) Věty 1. □

Lemma 7. *Nechť $\{(T_i, \leq) | i \in I\}$ je soubor všech maximálních stromů lokálně konečného lesa $(T, \leq)$. Následující podmínky jsou ekvivalentní:*

(i) $\mathrm{Max}(T, \leq) \cup \mathrm{Min}(T, \leq) = \emptyset$.

(ii) Pro libovolné $i \in I$ a každý prvek t z diskrétní stromové algebry $\mathcal{A}_i = (T_i, P_i)$ existují $\leq$ - srovnatelné prvky $x, y \in T_i, x \neq t \neq y$ tak, že $P_i(x, y, t) = t$.

(iii) Pro každé $i \in I$ a jakýkoli prvek $t \in T_i$ existují alespoň dva různé konce stromové algebry $(E(D_t), P^*), (E(D_t'), P^*)$ tak, že $D_t' \cup D_t$ je minimální LA-*retrakt stromu $(T_i, \leq)$.*

Důkaz. Implikace (i) $\Rightarrow$ (ii) bezprostředně vyplývá z definice diskrétní stromové algebry $\mathcal{A}_i$ realizované stromem $(T_i, \leq)$; podmínka (ii) říká jinými slovy, že jakýkoli prvek stromové algebry $\mathcal{A}_i$ je vnitřním prvkem vhodného ohraničeného intervalu, který je řetězcem vzhledem k uspořádání $\leq$.

(ii) $\Rightarrow$ (iii): Nechť $i \in I$ je libovolný index, $t \in T_i$ bude libovolný prvek. Podmínka (ii) znamená, že stupeň vrcholu t stromového grafu (T_i, ρ_i) je alespoň 2 a navíc $x \prec t \prec y$ pro vhodný zápis. Použitím matematické indukce konstruujeme řetězce

D_t, D_t' typu ω, ω^*, resp., přitom $D_t' \cup D_t$ je řetězec (s ohledem na $\leq$) typu $\omega^* + \omega$, tedy stromové algebry $(E(D_t), P^*), (E(D_t'), P^*)$ jsou různé. Pro každý prvek $x \in T_i$ označuje $f(x)$ jedinečný prvek $D_t' \cup D_t$ tak, že $\delta(x, f(x)) = 0$. Pak $f : T_i \to D_t' \cup D_t$ je LA-retrakt a $D_t' \cup D_t$ je evidentně minimální LA-retrakt stromu $(T_i, \leq)$.

(iii) $\Rightarrow$ (i): Tento důsledek dokazujeme pro horní les. Důkaz pro případ dolního lesa je podobný. Připusťme nejprve, že $\mathrm{Max}(T, \leq) \neq \emptyset$. Pro $a \in \mathrm{Max}(T, \leq)$ označme $(T_i, \leq)$ strom, jehož největší prvek je a. Podle bodu (iii) strom $(T_i, \leq)$ obsahuje nejméně dva různé nekonečné řetězce D_a, D_a' (typu) ω^*), z nichž každý je LA-retraktem $(T_i, \leq)$ (odpovídající retrakce jsou definovány podobně jako výše). $D_a' \cup D_a$ tedy nemůže být minimální retrakt.

Nyní připusťme, že $\mathrm{Max}(T, \leq) = 0$ a $\mathrm{Min}(T, \leq) \neq 0$. Nechť $b \in \mathrm{Min}(T, \leq)$ bude libovolný prvek a $i \in I$ index takový, že $b \in T_i$. Podle (iii) existují dva různé neohraničené intervaly D_b, D_b' algebry stromu $\mathcal{A}_i = (T_i, P_i)$ tak, že $D_b' \cup D_b$ je minimální LA-retrakce $(T_i, \leq)$. Pak $D_b' \cup D_b$ nemá horní hranici. Protože v opačném případě, pokud c označuje libovolnou horní hranici $D_b' \cup D_b$, pak jakýkoli lokální automorfismus g přiřazující k prvku $x > a$ prvek $g(x) \in D_b' \cup D_b$, není LA-retrakcí. Protože $D_b' \cup D_b$ je neohraničená (jako podmnožina $(T_i, \leq)$), přesně jeden z intervalů D_b, D_b' - řekněme D_b je řetězec typu ω. Protože $D_b \neq D_b'$, je množina $D_b \cap D_b'$ konečná. Označme d největší prvek $D_b \cap D_b'$ a položme $R = (D_b \div D_b') \cup \{d\}$ (kde $\div$ znamená symetrický rozdíl množin). Protože $(R, \leq)$ je řetězec typu $\omega^* + \omega$, jedná se o LA-retrakci $(T_i, \leq)$. Ale $R \subsetneq D_b' \cup D_b$, což je spor. Proto $\mathrm{Min}(T, \leq) \cup \mathrm{Max}(T, \leq) = \emptyset$. $\square$

Lemma 8. *Nechť* $(T, \leq)$ *je lokálně konečný les. Pak* $\mathrm{Max}(T, \leq) \cup \mathrm{Min}(T, \leq) = 0$ *právě tehdy, když pro každý maximální strom* $(T_i, \leq)$ *lesa* $(T, \leq)$ *grupoid* $(T_i, \circ)$ *je jednoduchý a* $\mathrm{End}(T_i, \circ) = \mathrm{LA}(T_i, \leq)$.

Důkaz. (Uvažujeme pouze případ horního lesa; v případě dolního lesa je důkaz analogický.)

Předpokládejme $\mathrm{Max}(T, \leq) \cup \mathrm{Min}(T, \leq) = \emptyset$. Nechť $(T_i, \leq)$ je libovolný maximální strom lesa $(T, \leq)$, $A \neq \emptyset$ je oboustranný ideál grupoidu $(T_i, \circ)$. Připusťme, že $T_i \setminus A \neq \emptyset$ a $t_0 \in T_i \setminus A$. Protože t_0 není minimální prvek $(T_i, \leq)$, existuje $t_1 \in T_i$, takový, že $t_1^+ = t_0$ a $t_1 \notin A$ (pro $t_1 \circ t_1 = t_0$). Protože $t \circ t = t^+$ pro každý prvek $t \in T$, máme, že A je duální ideál uspořádané množiny $(T_i, \leq)$, takže $t_0 < a$ pro některé $a \in A$. Potom $a \circ t_1 = t_1 \circ a = t_0 \notin A$, což je spor. Proto $A = T_i$.

Nyní dokážeme rovnost $\mathrm{LA}(T_i, \leq) = \mathrm{End}(T_i, \circ)$.
Předpokládejme $f \in \mathrm{LA}(T_i, \leq), a, b \in T_i$. Pokud $\delta(a, b) \leq 0$ máme $a \circ b = a^+$, $f(a \circ b) = f(a^+)$. Položme $c = a \vee b$. Potom $\mathrm{card}[f(a), f(c)] = \mathrm{card}[a, c] = \mathrm{card}[b, c] = \mathrm{card}[f(b), f(c)]$, proto $\delta(f(a), f(b)) \leq 0$ a odtud $f(a) \circ f(b) = (f(a))^+$. Protože f zobrazuje dvouprvkový řetězec $[a, a^+]$ izomorfně na dvouprvkový řetězec $[f(a), f(a^+)]$, máme $(f(a))^+ = f(a^+)$, takže $f(a \circ b) = f(a) \circ f(b)$, tj. $\mathrm{LA}(T_i, \leq) \subset \mathrm{End}(T_i, \circ)$.

Předpokládejme nyní, že $f \in \mathrm{End}(T_i, \circ)$, $a, b \in T_i, a < b$. Nechť $t_0 < t_1 < \cdots < t_n$ je takový řetězec, že $t_i \in T_i$, $t_0 = a$, $t_n = b, t_{i+1} = t_i^+$ pro $i = 0, 1, \ldots, n-1$. Potom $t_0 \circ t_0 = t_0^+ = t_1$, v důsledku toho $f(t_1) = f(t_0 \circ t_0) = f(t_0) \circ f(t_0) = (f(t_0))^+$. Odtud dostáváme, že interval $[a, b]$ stromu $(T_i, \leq)$ je izomorfní s intervalem $[f(a), f(b)]$, proto platí $f \in \mathrm{LA}(T_i, \leq)$, tj. $\mathrm{End}(T_i, \circ) \subset \mathrm{LA}(T_i, \leq)$ Proto $\mathrm{LA}(T_i, \leq) = \mathrm{End}(T_i, \circ)$.

Nyní připusťme $\mathrm{Min}(T_i, \leq) \neq \emptyset$ pro nějaký maximální podstrom lesa $(T, \leq)$. Předpokládejme $t_0 \in \mathrm{Min}(T_i, \leq)$ a položme $A = T_i \setminus \{t_0\}$. Pak pro každou dvojici prvků $t \in T$, $a \in A$ máme $t \circ a \in A$, $a \circ t \in A$, což znamená, že A je pravým ideálem grupoidu $(T_i, \circ)$, což je v rozporu s předpokladem. Proto $\mathrm{Min}(T_i, \leq) = \emptyset$. Připusťme, že strom $(T_i, \leq)$ má největší prvek s. Za předpokladu $\mathrm{LA}(T_i, \leq) = \mathrm{End}(T_i, \circ)$ a $s \circ s = s$, následně $f(s) = f(s \circ s) = f(s) \circ f(s)$, tj. $f(s) = s$ pro každý lokální automorfismus f stromu $(T_i, \leq)$. Protože $\mathrm{Min}(T_i, \leq) = \emptyset$, strom $(T_i, \leq)$ obsahuje klesající řetězec typu ω^* s největším prvkem $s, s = t_0 > t_1 > \ldots$ Položme $g(t_i) = t_{i+1}$ pro $i = 1, 2, \ldots,$ a $g(t) = g(t_i)$ pro jakýkoli prvek $t \in T_i$ s vlastností $\delta(t_i, t) = 0$. Pak $\mathrm{LA}(T_i, \leq)$, ale $g(s) = t_1 \neq s$, což je spor. Proto $\mathrm{Max}(T, \leq) \cup \mathrm{Min}(T, \leq) = \emptyset$. □

Lemma 9. *Nechť $(T, \leq)$ je lokálně konečný horní (dolní) les. Pokud $\mathrm{Max}(T, \leq) = \emptyset$ ($\mathrm{Min}(T, \leq) = \emptyset$), pak existuje transformace $f \in T^T$ s vlastností $C_T(f) = LA(T, \leq)$.*

Důkaz. Lokálně konečný horní les $(T, \leq)$ bez maximálních prvků je funkční graf nějaké transformace f množiny T (bez pevných bodů). Pro každý prvek $t \in T$ máme $f(t) = t^+$. Předpokládejme $g \in C_T(f)$, $t \in T$. Potom $(g(t))^+ = f(g(t)) = g(f(t)) = g(t^+)$. Pomocí matematické indukce získáme pro každou dvojici $t_1, t_2 \in T$, $t_1 < t_2$, že restrikce $g|[t_1, t_2]$ je pořádkový izomorfismus intervalu $[t_1, t_2]$ do intervalu $[g(t_1), g(t_2)]$. Pokud $g \in \mathrm{LA}(T, \leq)$, $t \in T$, pak $f(g(t)) = (g(t))^+ = g(t^+) = g(f(t))$, takže $g \in C_T(f)$. Proto máme $C_T(f) = \mathrm{LA}(T, \leq)$ v případě horního lesa.

Pokud $(T, \leq)$ je dolní les, pak každý z jeho prvků má nanejvýš jednoho předchůdce. Pro $t \in T \setminus \mathrm{Min}(T, \leq)$ položíme $f(t) = s$, kde $s^+ = t$; pokud $t \in \mathrm{Min}(T, \leq)$ položíme $f(t) = t$. Druhá část důkazu je podobná té předchozí. □

Poznámka 13. S ohledem na Lemma 9 dostaneme snadno následující tvrzení: Pokud $(T, \leq)$ je horní lokálně konečný les, $f : T \to T$ je zobrazení takové, že $f(t) = t^+$ pro každé $t \in T \setminus \mathrm{Max}(T, \leq)$ a $f(t) = t$ pro $t \in \mathrm{Max}(T, \leq)$ a $C_T(f) = \mathrm{LA}(T, \leq)$ právě tehdy, když $\mathrm{Max}(T, \leq) = \emptyset$.

Lemma 10. *Nechť $(T, \leq)$ je lokálně konečný strom bez maximálních a minimálních prvků. Pro každou dvojici různých prvků $a, b \in T$ existuje perfektní Alexandrovova topologie $\tau_{a,b}$ na T, takže $\mathrm{LA}(T, \leq) = S(T, \tau_{a,b})$ a body a, b jsou semiseparované v prostoru $(T, \tau_{a,b})$.*

Důkaz Lemmatu 10 najdeme v [58].

Poznámka 14. Zobecnění Lemmatu 10 pro lokálně konečný les je evidentní, protože z důkazu uvedeného lemmatu vyplývá, že není podstatné, zda je zvažováno zobrazení do sebe lokálně konečného stromu nebo zobrazení stromu do jiného stromu. (Všechny stromy by měly být bez maximálních a minimálních prvků.) Pojem strom v Lemma 10 tak může být nahrazen pojmem lesa a tvrzení zůstává pravdivé.

Lemma 11. *Nechť $(T, \leq)$ je lokálně konečný horní (dolní) antikořenný les, τ buď perfektní Alexandrovova topologie na T, takže $S(T, \tau) = \mathrm{LA}(T, \leq)$.*
Pak $\mathrm{Min}(T, \leq) = \emptyset$ ($\mathrm{Max}(T, \leq) = \emptyset$).

Důkaz. Předpokládejme, že $(T, \leq)$ je horní les. Protože $\mathrm{Max}(T, \leq) = \emptyset$, položením $f(t) = t^+$ pro libovolné $t \in T$, máme $C_T(f) = \mathrm{LA}(T, \leq)$ podle Poznámky 13. Nechť $\{(T_i, \leq) | i \in I\}$ je soubor všech maximálních podstromů lesa $(T, \leq)$. Protože $C_T(f) = S(T, \tau)$, podle [13] Věta 4.1 nastává přesně jeden z následujících případů:

1. $\mathrm{Min}(T_i, \leq) = \emptyset$ pro všechna $i \in I$.
2. $i \in I$ implikuje, že buď $(T_i, \leq)$ je řetězec typu $\omega^* + \omega$ nebo $T_i = K_i \cup \mathrm{Min}(T_i, \leq)$, kde $(K_i, \leq)$ je řetězec typu $\omega^* + \omega$ a $\mathrm{Min}(T_i, \leq) \neq \emptyset$.
3. $i \in I$ implikuje, že buď $(T_i, \leq)$ je řetězec typu $\omega^* + \omega$ nebo $T_i = \mathrm{Min}(T_i, \leq) \cup T_i'$, kde $\mathrm{Min}(T_i, \leq) \neq \emptyset$ ($\mathrm{card}\ T_i' = \aleph_0$) a pro každou dvojici prvků $s \in \mathrm{Min}(T_i, \leq)$, $t \in T_i'$ platí $s < t$.

S ohledem na výše uvedené podmínky *1, 2, 3* a rovnost $C_T(f) = S(T, \tau)$ snadno zjistíme, že všechny komponenty T_i, $i \in I$, jsou uzavřené podmnožiny (T, τ). Připusťme $\mathrm{Min}(T_i, \leq) \neq \emptyset$ pro nějaké $i \in I$. Nechť $t \in \mathrm{Min}(T_i, \leq)$ je libovolný prvek a položme $X = T_i \setminus \{t\}$. Pokud $t^+ = x^+$ pro některé $x \in X$, definujeme zobrazení $g : T \to T$ takto: $g(t) = x$, $g(s) = s$ pro $s \in T$, $s \neq t$. Je zřejmé, že $g \in C_T(f) = S(T_i, \tau)$ takže $\tau g(X) = g(\tau X) = X$, tedy $X = \tau g(X) = \tau \tau g(X) = \tau X$. Pokud prvek $x \neq t$ mající vlastnost $x^+ = t^+$ neexistuje (tj. $(T_i, \leq)$ je řetězec typu ω) $g : T \to T$ definujeme následovně: $g(s) = s^+$ pro $s \in T_i$, $g(s) = s$ pro $s \in T \setminus T_i$. Předpoklad $t \in \tau X$ implikuje (s ohledem na skutečnost $g \in S(T, \tau)$ a τ -uzavřenost T_i), že $\tau g(X) = g(\tau X) = X$, tj. $t \in \tau X = X$ - což je spor. Proto $t \notin \tau X$, tj. množina X je uzavřená. Pak $\{t\}$ je otevřená podmnožina, jejíž existence je v rozporu s předpokladem perfektnosti prostoru (T, τ). Následně množina $\mathrm{Min}(T_i, \leq)$ je prázdná.

Pokud $(T, \leq)$ je dolní les, definujeme transformaci f pomocí předchůdců prvků lesa $(T, \leq)$ a důkaz rovnosti $\mathrm{Max}(T, \leq) = \emptyset$ shoduje se s tím předchozím. □

Věta 18. *Jestliže $(T, \leq)$ je lokálně konečný antikořenný les, pak $S(T, \tau) = \mathrm{LA}(T, \leq)$ pro nějakou perfektní Alexandrovovou topologii τ na množině T právě tehdy, když existuje Alexandrovova topologie σ na T tak, že $\mathrm{LH}(T, \sigma) = \mathrm{LA}(T, \leq)$ a σ-uzávěr jakékoli neprázdné podmnožiny množiny T je nekonečný.*

Důkaz. Nechť τ je perfektní Alexandrovova topologie na T, takže $S(T,\tau) = \mathrm{LA}(T,\leq)$. Označme pomocí σ Alexandrovovu topologii duální k τ. Nejmenší σ-okolí jakéhokoli bodu $x \in T$ je τ-uzávěr $\tau\{x\}$, tedy $S(T,\tau) = \mathrm{LH}(T,\sigma)$. Protože $(T,\leq)$ neobsahuje žádný maximální nebo minimální prvek a každý lokální automorfismus $g \in \mathrm{LA}(T,\leq)$ takový, že g-obraz množiny T je řetězec typu $\omega^* + \omega$, je uzavřená deformace (T,τ), máme $\tau\{t\}$ je řetězec pro každé $t \in T$. Dále, $f^n \in S(T,\tau)$ pro každé $n \in \mathbb{N}$ (kde $f(t) = t^+$ pro $t \in T$), tedy $\tau\{t\}$ je kofinální podmnožina lesa $(T,\leq)$ v případě horního lesa a koiniciální podmnožina $(T,\leq)$ v případě dolního lesa. V důsledku toho σ -uzávěry singletonů jsou nekonečné. Nyní σ má být Alexandrovova topologie na T, která splňuje podmínky lemmatu. Potom σ^* (duální topologie k σ) má tuto vlastnost: $S(T,\sigma^*) = \mathrm{LA}(T,\leq^{-1}) = \mathrm{LA}(T,\leq)$ a prostor (T,σ^*) je perfektní. $\square$

Věta 19. *Nechť $(T,\leq)$ je lokálně konečný les. Pro jakoukoli dvojici různých prvků $a,b \in T$ existuje perfektní Alexandrovova topologie $\tau_{a,b}$, takže $S(T,\tau_{a,b}) = \mathrm{LA}(T,\leq)$ a body a,b jsou v prostoru $(T,\tau_{a,b})$ semiseparované právě tehdy, když pro jakoukoli dvojici různých prvků $c,d \in T$ existuje Alexandrovova topologie $\sigma_{c,d}$ na T, která má následující vlastnosti:*

(i) $\mathrm{LH}(T,\sigma_{c,d}) = \mathrm{LA}(T,\leq)$,

(ii) card $\sigma_{c,d}X \geq \aleph_0$ *pro každou neprázdnou podmnožinu* $X \subset T$,

(iii) body c,d jsou semiseparované v prostoru $(T,\sigma_{c,d})$.

Důkaz. Podobně jako v důkazu Věty 19 pomocí duální topologie dostáváme ekvivalenci výše uvedených podmínek. $\square$

Důkaz věty 15. Schéma důkazu je následující (1 - 5 jsou odpovídající podmínky):

$$\begin{array}{ccccc} 1 & \Longleftrightarrow & 2 & \Longrightarrow & 3 \\ & \nearrow\!\!\!\!\swarrow & \Uparrow & \swarrow\!\!\!\!\!\swarrow & \\ 5 & & 4 & & \end{array} \tag{3.11}$$

Ekvivalence $1 \iff 2$ je dána Lemmatem 6, implikace $2 \Longrightarrow 3 \Longrightarrow 4 \Longrightarrow 2$ Lemmatem 7 a ekvivalence $2 \iff 5$ je dána Lemmatem 8. $\square$

Důkaz Věty 16. vyplývá z věty 15 s ohledem na [25] věta 1, část (b). $\square$

Důkaz věty 17. Schéma implikací je následující:

$$\begin{array}{ccccc} 2 & & 4 & & \\ \Downarrow & \nwarrow & \Updownarrow & & \\ 1 & \Longrightarrow & 3 & \Longleftrightarrow & 5 \end{array} \tag{3.12}$$

Implikace $1 \implies 3$ vyplývá z Lemmatu 10 s ohledem na Poznámku 14 a Lemmatu 6. Implikace $3 \implies 2$ je triviální, $2 \implies 1$ je dána Lemmatem 11 s ohledem na Lemma 6. Ekvivalence $3 \iff 4$ je stanovena ve Větě 18 a ekvivalence $3 \iff 5$ ve Větě 19. □

Poznámka 15. Charakterizace obsažené ve Větě 17 byly získány pod základním předpokladem antikořennosti uvažovaných stromů. Charakterizace bez tohoto předpokladu nebo příležitostně s předpokladem existence větve ve všech stromech se zdají být otevřeným problémem. Další speciální řešení problémů nevyplývajících z výše uvedených úvah, je otázka realizace monoidů lokálních automorfismů konečných stromů uzavřenými deformacemi topologického prostoru.

Prezentovaná teorie lokálně konečných stromů a lesů s jejich transformacemi je aplikovatelná v oblasti struktur umělých neuronových sítí. Poznamenejme, že skutečné neuronové sítě jsou biologickými soustavami, protože jejich hlavní prvky, tedy neurony a synapse, jsou reálnými biologickými objekty. Prvky umělých neuronových sítí jsou naopak abstraktními entitami v podobě formalizované struktury, vlastností, procesů, projevů a chování neuronů a synapsí, takže umělé neuronové sítě lze považovat za systémy (ku příkladu v pojetí teorie systémů Masaroviče-Takahary), využívané k modelování chování neuronové sítě člověka a k řešení i nebiologických problémů. Neuron jakožto základní prvek nervové soustavy (neuronové sítě) člověka je tedy v intencích teorie systémů mikrobiologickou soustavou (systémem). Umělé neuronové sítě svou strukturou a procesy v nich probíhajících napodobují skutečné neuronové sítě, přičemž procesy napodobování jsou spojeny s procesy vytváření modelů. Pro umělou neuronovou síť je charakteristické, že do ní vstupuje vstupní signál $X = \{x_1, x_2 \ldots, x_m\}$ a vystupuje z ní signál výstupní $Y = \{y_1, y_2 \ldots, y_m\}$. Model funkce neuronové sítě je tedy jistou transformací vstupního signálu X na výstupní signál Y.

V rámci architektur neuronových sítí připomeňme - což bylo uvedeno již v úvodu práce - že vrstvené neuronové sítě se skládají z těchto entit[72]

a) vstupní vrstvy - má rozměr m_m (index udává počet prvků - neuronů), přičemž každý prvek dostává na svém vstupu pouze jednu složku x_i vstupního signálu,
b) skrytých vrstev - jejich počet označujeme r_s a rozměr s-té vrstvy značíme m_s,
c) výstupní vrstvy - rozměru m_n, přičemž z každého prvku vystupuje pouze jedna složka y_i.

Prvky skrytých vrstev jsou propojeny vzájemně i se vstupní a výstupní vrstvou způsobem každý s každým. V jednotlivých vrstvách neurony propojeny nejsou. Srovnej dopřednou neuronovou síť na obrázku 1.10, str. 22 v §1.2.2, včetně odkazu na publikaci [113] J. Šímy a R. Nerudy.

V těchto závěrečných úvahách stručně poukážeme na možnosti využití předcházejících teorií a naznačíme jejich návaznost na oblast neuronových sítí. Nejprve připomeňme práci Anny Jakovlevny Aizenštatové [4] o regulárních pologrupách endomorfismů uspořádaných množin. Podle věty 2.6 ze zmíněné práce (rovněž strana 68-69 této disertace) je monoid $\mathrm{End}(M, \leq)$ všech endomorfismů (tedy izotonních transformací do sebe) uspořádané množiny $(M, \leq)$, která není řetězcem ani protiřetězcem, regulární právě tehdy, když částečně uspořádaná množina $(M, \leq)$ má jeden ze čtyř tvarů uvedených ve větě 2.6 - mezi které také patří ordinální součet dvou protiřetězců a šestikoruna. Poznamenejme, že šestikorunou se rozumí šestiprvková uspořádaná množina $(M, \leq)$, $M = \{a_1, a_2, a_3, b_1, b_2, b_3\}$, v níž $a_1 < b_1$, $a_1 < b_2$, $a_2 < b_1$, $a_2 < b_3$, $a_3 < b_2$, $a_3 < b_3$ (přičemž $\{a_1, a_2, a_3\}$, $\{b_1, b_2, b_3\}$ tvoří protiřetězce). Podle věty 2.6 tedy uvažujeme-li fragmenty neuronové sítě tvořené vstupní vrstvou a skrytou vrstvou nebo skrytou vrstvou a výstupní vrstvou vrstevné neuronové sítě dostáváme uspořádané množiny neuronů jejichž monoidy endomorfismů jsou regulární. Rovněž šestikoruna neuronů je uspořádaná množina s regulárním monoidem endomorfismů.

Uveďme dále konkrétní příklad konečného podstromu výše zmíněné vrstevné neuronové sítě s regulárním monoidem lokálních automorfismů.

Buď $T = \{t_0, t_1, t_2, t_3, s_1, s_2\}$, přičemž uspořádání $\leq$ na stromu T je definováno takto: t_0 je kořen, tedy nejmenší prvek uspořádané množiny $(T, \leq)$, $t_0 < t_k$, $k = 1, 2, 3$, $t_3 < s_1$, $t_3 < s_2$, (trojice (t_0, t_1, t_2, t_3) a dvojice (s_1, s_2) tvoří protiřetězce, tedy t_1, t_2, s_1, s_2 jsou maximální prvky stromu $(T, \leq)$.) Nyní lze ověřit - např. výčtem, tedy sestrojením všech lokálních automorfismů stromu $(T, \leq)$, že monoid $\mathrm{LA}(T, \leq)$ obsahuje 36 prvků, včetně identického automorfismu $id_T : T \to T$. Dále lze ukázat, že každý z lokálních automorfismů $f \in \mathrm{LA}(T, \leq)$ je regulárním prvkem tohoto monoidu - ukážeme tento fakt na jednom příkladu: Označme

$$f = \begin{pmatrix} t_0 & t_1 & t_2 & t_3 & s_1 & s_2 \\ t_0 & t_1 & t_3 & t_3 & s_1 & s_2 \end{pmatrix}.$$

Zde f-obrazy uzlů (neuronů) v prvním řádku jsou uzly (neurony) umístěné pod nimi v druhém řádku této matice typu 2×6.

Definujeme zobrazení $g : T \to T$ předpisem

$$g = \begin{pmatrix} t_0 & t_1 & t_2 & t_3 & s_1 & s_2 \\ t_0 & t_1 & t_1 & t_3 & s_2 & s_2 \end{pmatrix}.$$

Potom platí

$$fgf(t_0) = fg(t_0) = t_0 = f(t_0),$$

$$fgf(t_1) = fg(t_2) = t_2 = f(t_1),$$

$$fgf(t_2) = fg(t_3) = f(t_2)(= f(t_3)),$$
$$fgf(t_3) = fg(t_3) = t_3 = f(t_3),$$
$$fgf(s_1) = fg(s_1) = s_1 = f(s_1),$$
$$fgf(s_2) = fg(s_1) = s_1 = f(s_2),$$

tedy $fgf = f$. Podobně lze ke každému zobrazení $f \in \mathrm{LA}(T, \leq)$ zkonstruovat zobrazení $g \in \mathrm{LA}(T, \leq)$ s výše uvedenou vlastností. Naznačeným postupem lze tedy ověřit, že monoid $\mathrm{LA}(T, \leq)$ je regulární.

Teorie, prezentované v předcházejících paragrafech umožňují podrobnější vnitřní analýzu stromových fragmentů vrstevných neuronových sítí a rovněž konstrukce transformací neuronových sítí kompatibilních s jejich strukturou. Dále pak graftingová (roubovací) aritmetika (což je jistá modifikace aritmetiky ordinální) poskytuje možnost dalších konstrukcí vertikálních propojování neuronových sítí. To vše skýtá náměty pro další studium a rozvíjení této aktuální problematiky v budoucnu.

Závěr

Disertační práce na téma Algebraizace a parametrizace přechodových relací mezi strukturovanými objekty s aplikacemi v oblasti neuronových sítí se zabývá parametrizací přechodových relací strukturovaných objektů v oblasti hypergrup a automatů s aplikací v oblasti neuronových sítí s vhodnou strukturou, aktivační funkcí, s cíleným zjednodušením modelu. Vzhledem k rozsahu oblasti neuronových sítí a také odlišnému způsobu fungování neuronových sítí včetně matematického podkladu bylo nezbytné nejprve vymezit základní oblast zkoumání neuronové sítě. Touto oblastí se stala dopředná neuronová síť - vícevrstevnatý perceptron (vícevrstvý perceptron – multilayer perceptron). Při následné analýze byla nalezena analogie s oblastí lineárních diferenciálních operátorů, která spolu s poznatky o parametrizovatelnosti relací mezi strukturovanými algebraickými objekty umožnila na bázi funkcionality jednotlivého neuronu konstruovat algebraické struktury umělých neuronů. Volba vah a výstupů, závislých na argumentu t umožnila vytvořit model časově proměnného neuronu, při omezené volbě přenosové funkce (identita), stejné pro celou neuronovou síť. Hlubší analogie umožnila vytvoření grupy a hypergrupy umělých neuronů, za využití kompatibility relací definovat uspořádání na těchto strukturách. Dalším postupem bylo definovat binární relaci a související uspořádání na množině umělých neuronů. Definice invariantní pologrupy a homomorfního zobrazení mezi strukturami umělých neuronů a algebraickými strukturami lineárních diferenciálních operátorů otevřela další možnosti popisu vlastností těchto struktur. Další fáze výzkumu spočívala v definici binární hyperoperace zobrazení z kartézského součinu grupy umělých neuronů do její potenční množiny a popisu vlastností transpoziční hypergrupy umělých neuronů. Volba sigmoidní přenosové funkce umožnila konstrukci P-hyperstruktur. Jako model, používající iteraci umělých neuronů byla použita konstrukce kaskády - akce aditivní grupy všech celých čísel na grupě umělých neuronů. Vektorový charakter dat v dopředné neuronové neuronové síti a definice struktury spojitých funkcí byly užity při rozšíření poznatků o řešitelnosti grup a hypergrup umělých neuronů. Grupové homomorfismy použité v kaskádě ukázaly další souvislosti mezi strukturami a hyperstrukturami umělých neuronů. Parciální pohled na části struktur dopředné neuronové sítě jako na lokálně konečné kořenové stromy a jejich algebraické struktury přinesl další pohled na modelování struktur a hyperstruktur umělých neuronů. Tyto poznatky otevírají další možnosti výzkumu neuronových sítí s užitím algebraických struktur včetně hyperstruktur.

Literatura

[1] Adams, D. P.: *Nomography: Theory and Application*, Archon Books, 1964.

[2] Adnaževič, D.: Some properties of A-spaces, *Soviet. Math. Dokl.* 1973, 14, 492–496.

[3] Adnaževič, D.: A-dopolněnija topologičeskich prostranstv,*Math. Balcanica* 1974, 4, 912.

[4] Aizenštat, A., J.: Reguljarnyje polugruppy endomorfizmov uporjadočennych množestv, *Učonnyje zapiski Leningrad.* Gos. Ped. Inst. 1968, 387, 3–11.

[5] Alexandrov, P. S.: Diskrete Räume, *Mat. Sb.* 1937, 2, 501–519.

[6] Amiri-Bideshki, M., Saeid, A. B., Ameri, R., Hošková-Mayerová, Š.: Prime filters of hyperlattices. *Analele Stiintifice Ale Universitatii Ovidius Constanta*, 2016, 24, 15–26.

[7] Bandelt, H. J.: Tolerance relations on lattices, *Bull. Austral. Math. Soc.* 1981, 23, 367–381.

[8] Bavel, Z. The source as a tool in automata, *Information and Control* 1971, 18, 140–155.

[9] Bédécarrats, A., Chen, S., Pearce, K., Cai, D., Glanzman, D. L.: RNA from Trained Aplysia Can Induce an Epigenetic Engram for Long-Term Sensitization in Untrained Aplysia, *eNeuro,* 2018, 5(3), 20182018.

[10] Behnke, S.: *Hierarchical Neural Networks for Image Interpretation*, Notes in Computer Science, Springer, Heidelberg, 2003.

[11] Beránek, J ., Chvalina, J.: Invariant subgroups of groups of ordinary second-order linear difierential operators, *Acta Math.* 2010, 13, 43–47.

[12] Birkhoff, G.: *Lattice Theory*, Amer. Math. Soc. Colloq. Publ., Providence 1979.

[13] Blažková, R., Chvalina, J.: Regularity and transitivity of local-automorphism semigroups of locally finite forest, *Arch. Math.* 1984, 20(4), 183–194.

[14] Blyth, T. S., Almeida, Santos, M. H. : Ordered regular semigroups with biggest associates, Discussiones Math., *General Algebra and Appl.*, 2019, 39(1), 5–21.

[15] Bishop, C. M.: *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

[16] Bishop, C. M.: Pattern recognition and feed-forward networks. In: Wilson R. A., Keil F. C. (eds) *The MIT encyclopedia of the cognitive sciences.* MIT, Cambridge, 1999, 629–632.

[17] Borůvka, O.: *Linear Differential Transformations of the Second Order.* English Universities Press, London, 1971.

[18] Borůvka, O.: Deformace a věty izomorfismu grup, In: *Základy teorie grupoidů a grup.* Nakladatelství Československé akademie věd, Praha, 1962.

[19] Borzooei, R. A., Varasteh, H. R., Hasankhani, A.: $\mathcal{F}$-Multiautomata on Join Spaces Induced by Differential Operators, *Applied Mathematics* 2014, 5, 1386–1391.

[20] Buchholz, S.: *A Theory of Neural Computation with Clifford-algebras*, Technical Report Number 0504, Institut für Informatik und Praktische Mathematik, Kiel, 2005.

[21] Clifford, A. H., Preston, G. B.: *The Algebraic Theory of Semigroups I.*, Amer. Math. Soc., Providence, 1977.

[22] Corsini, P.: Graphs and join spaces, *J. of Combinatorics, Information and System Science*, 1991, 4, 313–318.

[23] Corsini, P.: *Prolegomena of Hypergroup Theory.* Aviani Editore Tricesimo, 1993.

[24] Corsini, P., Leoreanu, V.: *Applications of Hyperstructure Theory.* Kluwer, Dortrecht, 2003.

[25] Čech, E.: *Topological Spaces* (revised by Z. Frolík and M. Katětov), Academia, Prague, 1966.

[26] Davvaz, B., Leoreanu Fotea, V.: *Hyperring Theory and Applications.* Hadronic Press, Palm Harbor, FL 2008.

[27] Dresher, M., Ore, O.: Theory of multigroups, *Amer. J. Math.*, 1938, 60, 705–733.

[28] Dörfler, W.: The cartesian composition of automata, *Math. Systems Theory*, 1978, 11, 239–257.

[29] Fahlman, S. E., Lebiere, C.: The cascade-correlation learning architecture. In: D. S. Touretzky (ed), *Advances in Neural Information Processing Systems*, 1990, 2, 524–532.

[30] Feit, W., Thomson, J. G.: A solvability criterion for finite groups and some consequences. *Proc. Nat. Acad. Sci.* 1962, 48, 968–970.

[31] Feit, W., Thomson, J. G.: Solvability of groups of odd order, *Pacific J. Math.*, 1963, 13, 775–1029.

[32] Goralčíková, A., Koubek, V.: Dominators and tree-semilattices, *Scripta Fac. Sci. Nat. Univ. Purkynianae Brunensis*, 1983, 13(8), 311–320.

[33] Groth, D., Skandier, T.: Network+ Study Guide, 4th ed, Sybex, 2005.

[34] Hagan, M., Demuth, H., Beale, M.: *Neural Network Design*, PWS Publishing, Boston, MA, 1996.

[35] Hall, M., Jr.: *The Theory of Groups.* Macmillan, New York, 1959.

[36] Hebb, D. O.: *The Organization of Behaviour: A Neuropsychological Theory.* Psychology Press, 2002.

[37] Hecht-Nielsen, R.: Counterpopagation networks. *Applied Optics,* 1987, 26(23), 4979–4984.

[38] Hickey, J. B.: Semigroups under a sandwich operation. *Proc. Edinburgh Math. Soc.* 1983, 26, 371–382.

[39] Hilbert, D.: Mathematische probleme. *Nachrichten der Akademie der Wissenschaften Göttingen,* 1987, 290–329.

[40] Hopfield, J. J., Tank, D. W.: Computing with neural circuits: a model. *Science*, 1986, 233, 625–633.

[41] Hošková-Mayerová, Š., Chvalina, J.: A survey of investigations of the Brno research group in the hyperstructure theory since the last AHA Congress. In: *AHA 2008: 10th International Congress - Algebraic Hyperstructures and Applications, Proceedings,* Brno: University of Defence, 2011, 71–83.

[42] Hošková-Mayerová, Š., Chvalina, J.: Discrete transformation hypergroups and transformation hypergroups with phase tolerance space. *Discrete mathematics*, 2008, 4133–4143.

[43] Hošková-Mayerová, Š., Chvalina, J.: Transposition hypergroups associated to linear partial differential operators. *Journal of Basic Science*, 2006, 1(3), 1–8.

[44] Hort, D., Chvalina, J.: On certain compatibility types of binary relations on hypergroups determined by quasi-orders. In: *Contributions to Gen. Algebra*, 1999, 11, 113–125.

[45] Chajda I.: *Algebraic Theory of Tolerance Relations*, UP, Olomouc, 1991.

[46] Chajda I.: Two constructions of compatible relations, *Czech. Math. J.*, 1978, 28, 439–444.

[47] Chajda, I., Zelinka, B.: Compatible relations on algebras, *Časopis pěst. mat.*, 1975, 100, 355–360.

[48] Chajda, I., Niederle, J., Zelinka, B.: On existence conditions for compatible tolerances, *Czech. Math. J.* 1976, 26, 304–311.

[49] Chvalina, J.: Commutative hypergroups in the sense of Marty and ordered sets, *Proceedings of the Summer School on General Algebra and Ordered Sets*, Olomouc, 1994, 19–30.

[50] Chvalina, J.: Characterizations of certain monounary algebras I. *Arch. Math.* 14(2), 1978, 85–98, II ibid. 14(3) 1978, 145–154.

[51] Chvalina, J.: *Functional Graphs, Quasi-ordered Sets and Commutative Hypergroups,* Vydavatelstvi Masarykovy univerzity, Brno, 1995.

[52] Chvalina, J.: Set transformations with centralizers formed by closed deformations of quasi-discrete topological spaces. In *Proc. IV. Prague Topological Symp. 1976,* Part B, Soc. Czech. Math. Phys., Prague, 1977.

[53] Chvalina, J., Hošková-Mayerová, Š.: General ω-hyperstructures and certain applications of those, *Ratio Math.*, 2012, 23, 55–72.

[54] Chvalina, J., Chvalinová, L.: Action of centralizer hypergroups of n-th order linear differential operators on rings on smooth functions. *Journal of Applied Mathematics*, 2008, 1(1), 45–53.

[55] Chvalina, J., Chvalinová, L.: Join spaces of linear ordinary differential operators of the second order. In: M. Bartušek, O. Došlý, (eds.) *Colloquium on Differential and Difference Equations, CDDE 2002, Proceedings*, Folia Fac. Sci. Nat. Univ. Masarykianae Brunensis, Mathematica 13, MU Brno 2003, 77–86.

[56] Chvalina, J., Chvalinová, L.: Locally finite rooted trees with regular monoids of local automorphismus. In: *Knižnice vědeckých a odborných spisů VUT v Brně.* B–119. Vysoké učení technické v Brně, Brno, 1988, 71–86.

[57] Chvalina, J., Chvalinová, L.: Modelling of join spaces by n-th order linear ordinary differential operators. In: *Fourth International Conference APLIMAT 2005.* Bratislava: Slovak University of Technology, Bratislava, 2005, 279–284.

[58] Chvalina, J., Chvalinová, L.: Transitively acting monoids of local automorphisms of localy finite trees, *Arch. Math.* 1983, 19(2), 71–82.

[59] Chvalina, J., Chvalinová, L.: State hypergroups of automata, *Acta Mathematica et Informatica Universitatis Ostraviensis*, 1996, 4, 105–120.

[60] Chvalina, J., Chvalinová, L., Moučka, J.: Solvability of a certain group of the third-order linear difierential operators. In: *Proc. XXX International Colloquium of the Management of Educational Process*, Univ. of Defence, Brno, 2012, 69–78.

[61] Chvalina, J., Novák, M., Staněk, D.: Sequences of groups and hypergroups of linear ordinary differential operators, *Ital. J. Pure Appl. Math.* (accepted for publication).

[62] Chvalina, J., Račková, P.: Join spaces of smooth functions and their actions on transposition hypergroups of second order linear differential operators, In: *Aplimat – Journal of Applied Math*, 2008, 1, 55–63.

[63] Chvalina, J., Račková, P.: Wallace-type compatibility of binary relations between products of hypergroups. In: *3. Mezinárodní Workshop.* FAST VUT v Brně, Brno, 2004, 1–4.

[64] Chvalina, J., Smetana, B.: Artificial neuron group and hypergroup actions. In: *Dynamical System Modelling and Stability Investigation*, Kyjev: Taras Shevchenko National University of Kyiv, Ukraine, 2019, 38–40.

[65] Chvalina, J., Smetana, B.: Compatible relations on state sets of quasi-automata. In:*Proceedings, 16th Conference on Aplied Mathematics Aplimat 2017.* Slovak University of Technology in Bratislava, 2017, 359–368.

[66] Chvalina, J., Smetana, B.: Groups and hypergroups of artificial neurons. In: *Proceedings, 17th Conference on Aplied Mathematics Aplimat 2018,* Slovak University of Technology in Bratislava, 2018, 232–243.

[67] Chvalina, J., Smetana, B.: Models of iterated artificial neurons. In *Proceedings, 18th Conference on Applied Mathematics APLIMAT 2019.* Slovak University of Technology in Bratislava, 2019, 203–212.

[68] Chvalina, J., Smetana, B.: Solvability of certain groups of time varying artificial neurons, *Ital. J. Pure Appl. Math.* (accepted for publication).

[69] Chvalina, J., Smetana, B.: Systems of fragments of artificial neural networks.. In*Proceedings, 19th Conference on Applied Mathematics Aplimat 2020.* Slovak University of Technology in Bratislava, 2020, 253–270.

[70] Chvalina, J., Svoboda, Z.: Sandwich semigroups of solutions of certain functional equations and hyperstuctures determined by sandwiches of functions. *Journal of Applied Mathematics*, 2009, II(1), 35–43.

[71] Ivakhnenko, A. G.: Polynomial theory of complex systems. *IEEE Transactions on Systems, Man, and Cybernetics 1*, 1971, 4, 364–378.

[72] Janíček, P.: *Systémové pojetí 1. vybraných oborů pro techniky hledání souvislostí.* VUTIUM, Brno, 2007.

[73] Junqueira, L. C. U., Carneiro, J., Kelley, R. O.: *Basic Histology.* Appleton and Lange, Norwalk, 1999.

[74] Klein, E., Thompson, A. C.: *Theory of Correspondences*, Wiley, New York, 1984.

[75] Koukolík, F.: *Lidský mozek: funkční systémy, norma a poruchy.* 3. vyd., Galén, Praha, 2012.

[76] Koskela, T.: *Neural Network Methods in Analysing and Modelling Time Varying Processes* (Disertační práce). Helsinki University of Technology, Helsinki, 2003.

[77] Laufberger, V. *Vzruchová theorie: učebnice fysiologie jednání na základě nové theorie paměti.* Spolek českých lékařů, Praha, 1947.

[78] Leoreanu-Fotea, V., Ciurea, C. D.: On a P-hypergroup. *Journal of Basic Science*, 2008, 4(1), 75–79.

[79] Lippmann, R. P.: An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine,* 1987, 1–22.

[80] Lorentz, G. G.: *Approximation of Function.* Halt, Reinhart, and Winston, New York, 1966.

[81] Magill, J. K. D., Subbiah, S.: Green relations for regular elements of semigroups of endomorphisms, *Canad. J. Math.*, 1974, 6, 1484–1497.

[82] Marty, F. Sur une généralisation de la notion de groupe. *IV Congrès des Mathématiciens Scandinaves*, 1934, 45–49.

[83] Maxson, C. J.: Semigroups of order-preserving partial endomorphisms on trees I, *Coll. Math.* 1974, 23, 25–37.

[84] Minsky, M. L., Papert, S. A.: *Perceptrons.* Cambridge, MA: MIT, 1969. Expanded Edition 1990.

[85] McCulloch, W., Pitts, W.: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 1943, 5, 115–133.

[86] Molnár, Z.: Úvod do problematiky umělých neuronových sítí. [online] *Elektrorevue*, 2000, 13, [cit. 22. 01. 2020]. Dostupné z URL: `http://www.elektrorevue.cz/clanky/oldindex.html`

[87] Munkres, J. R.: *Topology.* Prentice Hall, 1999.

[88] Mushtaq, Q., Khan, M.: Ideals in AG-band and AG -groupoid, *Quasigroups and Related Systems*, 2006, 14, 207–215

[89] Nebeský, L.: *Algebraic Properties of Trees*, Acta Univ. Carolinae Philologica, Praha, 1969.

[90] Nelson, E.: Homomorphisms of mono-unary algebras, *Pacific J. Math.* 1982, 99, 427–429.

[91] Neuman, F.: From local to global investigations of linear differential equations of the n-th order, *Jahrbuch Uberblicke Mathematik* 1984, 55–80.

[92] Neuman, F.: *Global Properties of Linear Ordinary Differential Equations,* Academia Praha - Kluwer, Dortrecht, 1991.

[93] Neuman, F.: Global theory of ordinary linear homogenerous differential equations in the real domain. *Aequationes Math.* 1987, 33, 123–149.

[94] Novák, M.: n-ary hyperstructures constructed from binary quasi-ordered semigroups. *Analele Stiintifice Ale Universitatii Ovidius Constanta, Seria Matematica*, 2014, 22, 147–168.

[95] Novák, M.: Some basic properties of EL-hyperstructures. *European Journal of Combinatorics*, 2013, 34(2), 446–459.

[96] Novák, M.: On EL-semihypergroups. *European Journal of Combinatorics*, 2015, 6(44), 274–286.

[97] Novák, M., Cristea, I. Composition in *EL*–hyperstructures. *Hacet. J. Math. Stat.* 2019, 48, 45–58.

[98] Novotný, M.: Sur un problème de la thèorie des applications, *Publ. Fac. Sci. Univ. Masaryk*, 1953, 344, 53–64.

[99] Novotný, M.: Über Abbildungen von Mengen, *Pacific. J. Math.* 1963, 13, 1347–1359.

[100] Novotný, M.: Characterizations of the category of connected machines, *Found. Control. Engineering 7*, 1982, 3, 169–179.

[101] Novotný, M.: Zaměnitelnost endomorfismů lineárních prostorů, *Čas. pěst. mat.* 1982, 107, 124–138.

[102] Prenowitz, W., Jantosciak, J .: Geometries and join spaces. *J. Reine Angew. Math,* 1972, 257, 100–128.

[103] Procházka, L., Bican, L., Kepka, T., Němec, P.: *Algebra.* Academia, Praha, 1990.

[104] Rachůnek, J.: Quasi-orders of algebras, *Časopis pěst. mat.* 1979, 104, 327–337.

[105] Rival, I.: *The retract construction, Ordered Sets,* In: I. Rival (ed.), *Ordered Stes*, Dortrecht Reidel, Dortrech, 1982, 97–122.

[106] Rosenblatt, F.: *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* Spartan, Washington DC, 1962.

[107] Rosenblatt, F.: *The Perceptron, A Perceiving and Recognizing Automaton,* Project Para, Report No. 85-460-1, Cornell Aeronautical Laboratory, New York, 1957.

[108] Rumelhart, D. E., Hinton, G. E., Williams, R. J.: Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Micro structure of Cognition,* MIT Press, Cambridge, MA, 1986, 318–362. (Reprinted in Anderson and Rosenfeld (1988))

[109] Skornjakov, L. A.: Unary algebras with regular endomorphism monoids, *Acta. Sci. Math.* 1978, 40, 375–381.

[110] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learning Res.* 2014, 15, 1929–1958.

[111] Šteinbuk, V. B.: Polugruppy endomorfizmov sistěm nasyščennych cepěj, *Sovremennaja algebra*, 1980, 150–159.

[112] Stoll, R. R.: *Sets, Logic and Axiomatic Theories.* W. H. Freeman and Company, 1974.

[113] Šíma, J., Neruda, R.: *Teoretické otázky neuronových sítí.* MATFYZPRESS, Praha, 1996.

[114] Volná, E.: *Neuronové sítě 1.* 2. vyd., Ostravská univerzita, Ostrava, 2008.

[115] Vougiouklis, T.: Cyclicity in a special class of hypergroups, *Acta Universitatis Carolinae. Mathematica et Physica,* 1981, 22, 3–6.

[116] Vougiouklis, T.: Generalization of P-hypergroups, *Rendiconti del Circolo Matematico di Palermo*, 1987, 36, 114–121.

[117] Vougiouklis, T., Konguetsof, L.: P-hypergoups. *Acta Universitatis Carolinae. Mathematica et Physica*, 1987, 28, 15–20.

[118] Warner, M. W.: Semigroup, group quotient and homogeneous automata, *Inform. Control* 1980, 47, 59–66.

[119] Wells, C.: Centralizers of transitive semigroup actions and endomorphisms of trees, *Pacific J. Math.* 1976, 64, 265–271.

[120] Widrow, B.: *An Adaptive Adaline Neuron Using Chemical Memistors,* Stanford Electronics Laboratories Technical Report 1553-2, October 1960.

[121] Zelinka, B.: Infinite tree algebras, *Čas. pěst. mat.* 1982, 107, 59–68.

[122] Zhao, Y., Deng, B., Wang, Z.: Analysis and study of perceptron to solve XOR problem, In: *Proc. 2th Int. Workshop Auto. Decentralized Syst.,* 2002, 168–173.

Seznam symbolů, veličin a zkratek

ADALINE adaptivní lineární neuron – ADAptive LInear NEuron
CIFAR Kanadský institut pokročilého výzkumu – Canadian Institute for Advanced Research
CN Cliffordův neuron – Clifford neuron
DARPA Agentura ministerstva obrany pro pokročilé výzkumné projekty – Defense Advanced Research Projects Agency
DSP číslicové zpracování signálů – Digital Signal Processing
GMAC Generalizovaná podmínka smíšené asociativity – Generalized Mixed Associativity Condition
IEEE Institut pro elektrotechnické a elektronické inženýrství – Institute of Electrical and Electronics Engineers
INNS Mezinárodní společnost neuronových sítí – International Neural Network Society
GMDH Skupinová metoda zpracování dat – Group method of data handling
MAC Podmínka smíšené asociativity – Mixed Associativity Condition
MATLAB počitačový program pro práci s maticemi, grafy, neuronovou sítí... – matrix laboratory
MLP vícevrstvý perceptron – multilayer perceptron
PDP Skupina paralelního distribuovaného zpracování – Parallel Distributed Processing Group
poset částečně uspořádaná množina – partially ordered set
SP substituční vlastnost – substitution property
SP1 substituční vlastnost prvního typu – substitution property of the first type
SP2 substituční vlastnost druhého typu – substitution property of the second type
SP3 substituční vlastnost třetího typu – substitution property of the third type
SVM metoda podpůrných vektorů – Support vector machines
W-kompatibilita W-kompatibilita – kompatibilita Wallaceho typu
XOR Exkluzivní disjunkce