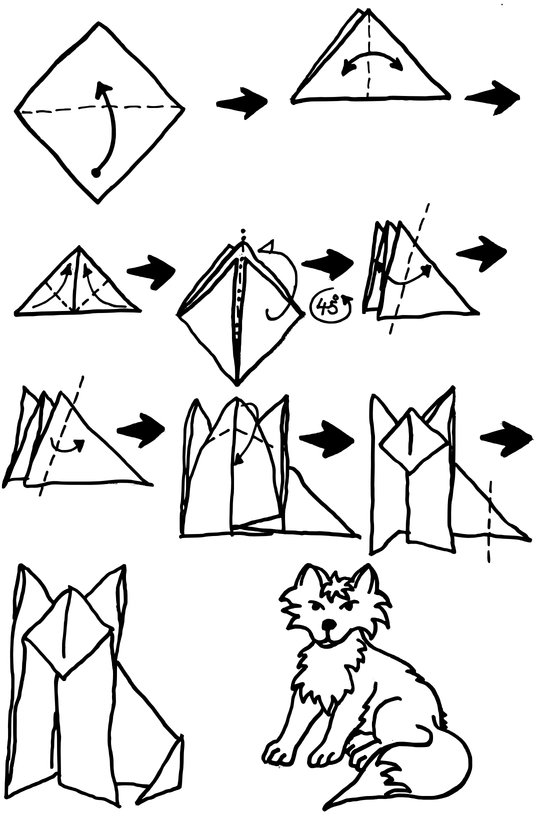
Studentský matematicko-fyzikální časopis

Ročník XIX / Číslo 2



Časopis M&M a stejnojmenný korespondenční seminář je určen pro studenty středních škol, kteří se zajímají o matematiku, fyziku či informatiku. Během školního roku dostávají řešitelé zdarma čísla se zadáním úloh a témat k přemýšlení. Svá řešení odesílají k nám do redakce. My jejich příspěvky opravíme, obodujeme a pošleme zpět. Nejzajímavější řešení otiskujeme.

Milí řešitelé,

přinášíme vám druhé číslo našeho časopisu a s ním zadání nových úloh a témat. Další témata už letos zadávat nejspíš nebudeme. Doufáme, že v naší současné nabídce naleznete něco, co vás zaujme. Nově se můžete ponořit do tajů oboru matematiky honosně nazývaného teorie čísel nebo zapřemýšlet nad podobností úvah v různých odvětvích fyziky.

Na konci čísla naleznete letošní první organizátorský článek, ve kterém vám (R)adim představí neuronové sítě. Dopředu můžeme prozradit, že v dalším čísle se můžete těšit i na článek od jednoho z řešitelů.

Všechny případné zájemce o studium na MFF UK bychom chtěli pozvat na Den otevřených dveří naší fakulty, který se koná ve čtvrtek **29. listopadu** v Praze. Můžete se dozvědět, jak studium na MFF vlastně probíhá, a prohlédnout si různá pracoviště. Více informací a především podrobný program naleznete na adrese `http://www.mff.cuni.cz/verejnost/dod/`.

organizátoři M&M

Zadání úloh

Termín odeslání druhé série: 10. 12. 2012

Krajina hnědá je a často mrholí,
jen vůni hlíny teď ucítíš na poli.
Po cestě válí se ztracené brambory,
šedý mrak bouřkový valí se nad hory.

Úloha 2.1 – Nad Tatrou se blýská (5b)

Několik kamarádů se rozhodlo, že změří rychlost zvuku ve vzduchu. Počkali na bouřku, vyběhli na kopce (každý na jeden) a čekali na blesky. Ke každému blesku si všichni zaznamenali časový odstup mezi zábleskem a zazněním hromu. Dokážou z těchto údajů spočítat rychlost zvuku (a jak), pokud

(i) jsou dva a znají svou vzdálenost,
(ii) jsou tři a znají všechny svoje vzdálenosti,
(iii) jsou tři a znají jen jednu vzdálenost,
(iv) jsou tři a neznají ani jednu vzdálenost?

Měkce a tlumeně zní Tvoje kroky,
na bílém poprašku zůstanou stopy.
Děti si zakřičí: „Pozor, já jedu!“,
jsou celé promrzlé jak kostka ledu.

Úloha 2.2 – Kostka (4b)

Riki přemýšlel, jak by vypadala hra Člověče, nezlob se s kostkami-nekostkami o všelijakém počtu stěn, řekněme 2 až 20. Pro tyto kostky klasicky platí, že po hození nejvyššího čísla háže hráč znovu a hody se sčítají. Pravděpodobnost padnutí všech čísel je stejná. Otázkou je, kolikastěnná kostka je nejvýhodnější, tedy dosáhne v této hře nejvyšší střední hodnoty součtu hodů.

Stomy se obsypou bílými květy,
lesy se stávají vodními světy.
Sluníčko teple zas na zemi svítí
v trávě se objeví prvních pár kvítí.

Úloha 2.3 – Kvítky (3b)

Riki si natrhal 9999 kvítků – některé jsou bílé a některé žluté. Když je rovnal doma k usušení, všiml si, že jsou zajímavě seřazené: žluté jsou právě ty kvítky, jejichž pořadí lze zapsat jako $\lfloor 2x \rfloor + \lfloor 4x \rfloor + \lfloor 8x \rfloor + \lfloor 16x \rfloor$.[1] Kolik kvítků je bílých?

Zahrady plné jsou lákavých plodů,
dáš si v ten horký den nanuk, či sodu?
Když kropíš hadicí, vidět je duhu,
děti si v rybníce plavají v kruhu.

Úloha 2.4 – Kruhová chodba (2b)

Kosmonaut je v dlouhé kruhové chodbě, která vede po obvodu neznámé kosmické lodi. V celé chodbě jsou na stěnách v pravidelných intervalech rozmístěny žárovky. Ty jsou bez zjevného řádu rozsvícené, nebo zhasnuté a u každé je příslušný vypínač. Kosmonaut by chtěl obejít celou chodbu a poznat, že ji už obešel. Může si pro svou potřebu přepínat žárovky, jinak je ale chodba všude stejná. Jak to má provést, aby se co nejmíň nachodil a nemusel si toho moc pamatovat?

[1] Pro nějaké reálné x, $\lfloor x \rfloor$ je dolní celá část čísla x, tzn. největší celé číslo, které je menší rovno x

Zadání témat

Téma 4 – Věty o čtvercích

Jedním z klasických problémů teorie čísel je otázka, která přirozená čísla lze zapsat jakou součet čtverců (druhých mocnin) několika celých čísel. Tento problém je již docela důkladně prozkoumán. Asi zásadní je následující na první pohled překvapivá věta.

Věta (o čtyřech čtvercích, Lagrange, 1770): Každé přirozené číslo n lze zapsat ve tvaru

$$n = a^2 + b^2 + c^2 + d^2,$$

kde a, b, c, d jsou celá čísla.

Její důkaz není jednoduchý. Pokud by vás přesto zajímal, můžeme se podívat na konferenční příspěvek Prof.[MM] Štěpána Šimsy v čísle 6–7 16. ročníku, kde je podrobně popsán. Jistě vás napadne otázka, jestli jsou čtyři čtverce opravdu potřeba. Odpověď zní, že jsou. Existují totiž čísla, která pomocí tří čtverců zapsat neumíme.

Věta (o třech čtvercích, Gauss): Přirozené číslo n lze zapsat ve tvaru

$$n = a^2 + b^2 + c^2,$$

kde a, b, c jsou celá čísla, právě tehdy, když $m \neq 4^m(8k+7)$, k, $m \in \mathbb{N}$.

Podobně jsou důkladně prozkoumána i čísla, která lze zapsat jako součet dvou čtverců. Tímto problémem se zabýval další významný matematik – Pierre Fermat. Prvním krokem v důkazu věty o dvou čtvercích je tvrzení, která prvočísla lze zapsat jakou součet dvou čtverců. Zkuste se nad tím zamyslet. Není to až tak obtížné.

Nabízí se i další otázka. Pokud máme dvě čísla zapsána jakou součet dvou čtverců, lze tak zapsat i jejich součet? A jejich součin?

Čísla, která lze zapsat jakou součet dvou čtverců, již byla popsána. Jak je to ale s rozdílem? Která celá čísla lze zapsat jako rozdíl dvou čtverců?

Pokud chcete náročnější námět k přemýšlení, zkuste dokázat, že každé přirozené číslo $m > 1$ lze zapsat ve tvaru

$$m = a^2 + b^2 + c^2 + 2^k,$$

kde a, b, c, $k \geq 0$ jsou přirozená. Při dokazování doporučuji využít větu o třech čtvercích. Dokonce platí i silnější tvrzení. A to, že každé přirozené $m > 1$ lze zapsat ve tvaru

$$m = a^2 + b^2 + c^2 + 2^{2k}.$$

To je zajímavé v tom, že nám větu o třech čtvercích převádí na větu o čtyřech čtvercích. Zkuste se zamyslet i nad ním. Hodně štěstí při bádání.

Kuba

Téma 5 – Analogie s elektrickými obvody

Nejdříve bychom si měli vysvětlit, co to je ona analogie. Snahou fyziky je nalézt co nejjednodušší a nejpřesnější modely, které popisují okolní svět. V různých případech se můžeme setkat s modely, které sice popisují odlišné fyzikální děje, ale způsob, jakým to dělají, je podobný, tedy analogický. Analogie, které budeme hledat my, se pojí k elektrickým obvodům.

U elektrických obvodů obvykle vyšetřujeme, jaké proudy a napětí nám vznikají v jednotlivých částech obvodu. Představujeme si, že u plus[2] kontaktu baterky máme kuličky, kterým říkáme elektrony, a ty se snaží dostat na druhou stranu, kde je naopak elektronů nedostatek. Cesta, kterou se můžou pohybovat, vznikne propojením obvodu pomocí vodičů. Pracujeme s pojmy napětí, proud, vodivost, ...

Obdobně se ovšem můžeme dívat například na tekutiny, tedy kapaliny a plyny. I ony mají tendence proudit odněkud někam. V případě plynů se snaží dostat z místa, kde má plyn vysoký tlak, do místa, kde je tlak menší. Jak možná tušíte, dostáváme zde analogii mezi tlakem a napětím. Analogie mezi proudem plynu a proudem elektronů je nasnadě. Dále si uvědomíme, že můžeme zavést pojem vodivosti potrubí, neboť přes potrubí s malým průřezem nám bude plyn téct jen těžko a přes trubku s velkým průřezem nám poteče jedna radost. Nyní můžeme sestavit ekvivalent k Ohmově zákonu. A pak můžeme postupovat dále a ukázat, že i pro plyny platí Kirchhoffovy zákony.

Co od vás tedy očekáváme? Pokuste se vymyslet analogii mezi prouděním plynu či kapaliny a prouděním elektronů ve vodičích. Nejdříve si ujasněte, co je plynové či tekutinové napětí, proud, vodivost, kapacita... Pak zkuste sestavit základní zákony jaké známe z elektřiny, tedy Ohmův zákon, Kirchhoffovy zákony, či jiné. Hlavně se ale zamyslete, zda tyto úvahy dávají smysl. Zkuste se pustit do vymýšlení, jak by mohl vypadat takový plynový či tekutinový kondenzátor, či co by v našem modelu mohlo představovat magnetické pole. Můžeme si pak vymyslet ekvivalent cívky, či jiných složitějších elektrických součástek. Pokud máme k dispozici cívku a kondenzátor, tak proč se nepustit do úvah nad střídavými obvody? A mimochodem, jak může vypadat náš plynový či tekutinový „ampérmetr“?

Fantazii se u tématka meze nekladou. Navíc správné řešení u tohoto tématka neexistuje, je možné, že se vám podaří vymyslet více funkčích modelů. Proto budeme velice rádi, pokud nám zašlete i dílčí myšlenky, či budete ve svých komentářích reagovat na články svých kolegů.

(R)adim

[2] Případně mínus, záleží na úhlu pohledu. :-)

Řešení témat

Téma 1 – Stavba století

Přinášíme vám čerstvé zprávy ze staveniště. V prvním kole předložily svoje návrhy tři odvážné stavební společnosti a bylo jim uděleno stavební povolení. Tyto firmy jsou vedené Patrikem Nácovským, Jiřinou Svobodovou a Mgr.$^{\mathrm{MM}}$ Pepou Svobodou. Do druhého kola výběrového řízení následně dorazily další žádosti (příspěvky). Financování bylo přiznáno Filipu Homzovi, Pavlu Součkovi a Arance Hruškové.

Většina návrhů tvrdí:

1. V nekonečné ulici nelze postavit věž větší než 3.
2. Na pozemku 3×3 lze postavit nanejvýš věž velikosti 4.
3. Pokud dostanou parcelu aspoň 5×5, tak postaví věž velikosti 5.

Ukázali, jak takovéto věže postavit a předložili důkazy, které říkají, že větší věž nepostaví. Je to opravdu největší možná výška nebo jsou v jejich důkazech chyby?

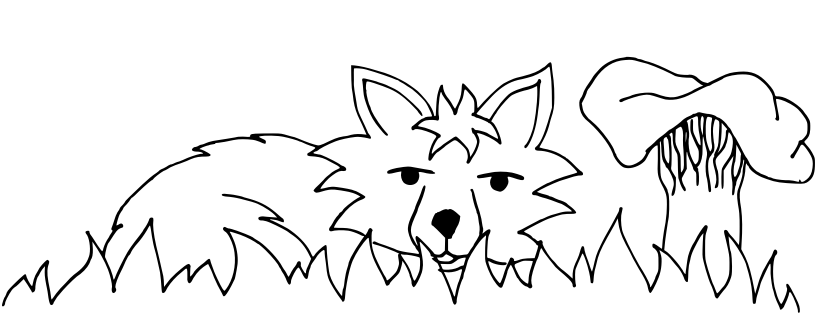
Podařilo se nám sehnat větší rozpočet a proto bude oslava mnohem větší, než se předpokládalo. Stavět (slavit) se bude i ve vícedimenzionálních prostorech, nejen na dvojrozměrných (čtvercových) parcelách.

Zkuste popsat, jak by to vypadalo ve čtyřrozměrném prostoru. Pokuste se odpovědět na otázky z minulého čísla a dále celý problém rozvinout do více dimenzí. A jak by to vypadalo, kdybychom dostali parcelu, která by měla n dimenzí?

Protože práce je stále dost, tak by nás zajímalo, jak velký pozemek budeme potřebovat, když budeme chtít postavit více věží, které budou stejně velké? Dejte si pozor, aby se vám nestalo, že postavíte tři věže stejné velikosti vedle sebe, to by se pak z nich stala věž větší. Zkuste to nějak (např. na postavení a věží bude určitě stačit pozemek $M \times N$ a na pozemku $m \times n$ se a věží určitě nepostaví) obecně vyjádřit.

Samozřejmě se můžete zabývat i jinými problémy, než které tu jsou nastíněny. Projektování zdar!

xlfd



Neuronové sítě

S pojmem neuronových sítí máme spojenu spíše biologii či lidský mozek. Přesto jsou neuronové sítě mocným nástrojem při řešení nejrůznějších informatických, fyzikálních či jiných problémů. A to teď nemluvíme o neuronových sítích, které má každý z nás uloženy v lebce, ale o umělých sítích, které si můžeme vytvořit v počítači.

Počítačové neuronové sítě, jak již název napovídá, mají svou předlohu v přírodě. Konkrétně v lidském mozku. Neuron je nervová buňka, která je schopna přijmout, zpracovat a vyslat signály. Neurony, o kterých budeme mluvit, si můžete představit jako krabičku, do které vedou přívody a ze které vede jeden vývod. Přičemž vývody můžou být zapojeny do přívodů některých jiných neuronů.

Signály, které se nám šíří mezi neurony, můžeme reprezentovat reálnými čísly v intervalu mezi nulou a jedničkou včetně. Máme-li neuron se vstupy x_1, x_2, …, x_n, tak jeho výstup y můžeme popsat vztahem

$$y(x_1, x_2, \ldots, x_n) = f\left(\left[\sum_{i=0}^{n} w_i x_i\right] - p\right),$$

kde w_i jsou vstupní váhy neuronu, p je práh neuronu a $f()$ značí tzv. přenosovou funkci neuronu. Prozatím si ji můžeme představit tak, že $f(x) = 1$ pro $x > 0$ a $f(x) = 0$ pro $x \leq 0$. Tedy pokud vstupní signály vynásobené patřičnými váhami jsou v součtu větší než práh neuronu, tak se neuron sepne, jinak se nesepne.

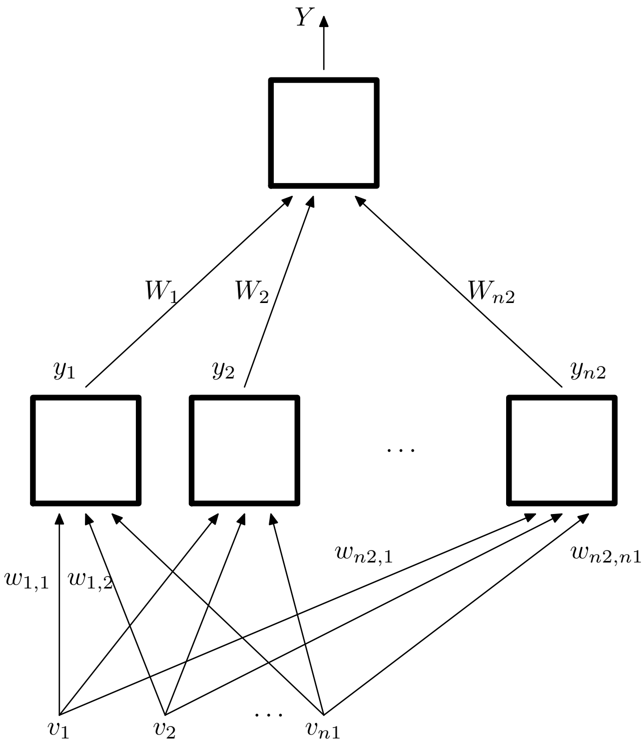
Pokud bychom si vzali jednoduchý příklad, kdy máme neuron se dvěma vstupy x_1 a x_2 a prahem $p = 1$, tak například pro váhy $w_1 = w_2 = 0{,}75$ máme neuron, který představuje logickou funkci AND. Neboli, pouze pokud $x_1 = x_2 = 1$, tak dostaneme na výstupu $y = 1$, jinak $y = 0$. Pokud váhy změníme například na $w_1 = w_2 = 1{,}5$, tak při zachování prahu $p = 1$ dostaneme logickou funkci OR.

Při práci s neuronovou sítí postupujeme tak, že si zvolíme problém, který má síť řešit. Ten si vyjádříme pomocí funkce $F(v_1, v_2, \ldots, v_n)$, kde v_1, v_2, …, v_n jsou vstupní proměnné našeho problému. Můžeme například chtít, aby naše neuronová síť uměla rozpoznávat obraz. Vstup budeme získávat z černobílé kamery a můžeme jej reprezentovat jako jednotlivé pixely, které očíslujeme v_1, v_2, …, v_n. Velikost v_i určuje stupeň šedi. Výstup neuronové sítě může vypadat například tak, že pokud kamera snímá domeček, který může být mírně zašuměný, tak jeden výstup bude 1, zbytek nula, pokud kamera snímá autíčko, tak druhý výstup bude 1 a zbytek nula…

Aby toto naše síť uměla, tak ji to nejdříve musíme naučit. To uděláme tak, že náhodně zvolíme váhy a prahy neuronů a budeme síti předkládat učební obrázky a podle toho, jak bude odpovídat, jí budeme měnit váhy a prahy neuronů až do chvíle, kdy se sítí budeme spokojeni.

Síť, která bude schopná řešit nějaký podobný úkol, bude muset být poměrně složitá. Pokud by obsahovala málo neuronů, třeba jeden, tak bychom neuměli nastavit váhy w tak, abychom dostali požadovaný výsledek.

Ukážeme si jednoduchý příklad, kdy síť sestavíme tak, že bude mít tzv. skrytou vrstvu neuronů a jeden výstupní neuron. Situace je znázorněna na obrázku c1.1.



Obr. c1.1 – Příklad neuronové sítě.

Pro skrytou vrstvu platí

$$y_k = f\left(\left[\sum_{i=0}^{n_1} w_{k,i} v_i\right] - p_k\right), k \leq n_2 ,$$

a pro výstupní neuron platí

$$Y = f\left(\left[\sum_{i=0}^{n_2} W_i y_i\right] - P\right) .$$

Na problém, který řešíme, se můžeme také dívat tak, že máme nějakou chybovou funkci

$$E = \sum_s \frac{1}{2}\left(Y(s) - V(s)\right)^2 ,$$

která nám říká, jak moc naše neuronová síť chybuje. Přičemž sumy přes s provádíme přes všechny zadané výsledky, $V(s)$ značí správné výsledky. Jedná se o zkušební sadu, kterou předkládáme síti. U této funkce hledáme minimum. Funkce z definice nemůže být menší než nula, najdeme-li tedy opravdu minimum, našli jsme nejlepší možnou síť. Musíme si ovšem dát pozor, abychom nedoputovali pouze do lokálního minima. To si můžeme představit tak, že dojdeme do nějakého uzavřeného údolí, ovšem pokud bychom vylezli ještě na jeden kopec a pak sešli dolů, tak bychom našli lepší údolí.

Hledání minima je trochu složitější. Nejsnáze se to provádí pomocí derivací. Vysvětlení derivací dalece přesahuje rozsah tohoto textu, proto zde odvození nebudeme uvádět. Znalí zájemci si jej můžou provést sami. Funkce $E()$ závisí na $y()$, což je funkce obsahující $f()$. Ovšem chceme-li $E()$ derivovat, tak potřebujeme, aby používané funkce neměly skoky, což $f()$ rozhodně nesplňuje. Proto ji nahradíme funkcí

$$f(\nu) = \frac{1}{1 + \exp(-\lambda\nu)}\,,$$

kde λ značí rychlost přeběhu této funkce. Pro $\lambda \to \infty$ dostaneme naši skokovou funkci. Tuto funkci už umíme derivovat. Je možné používat i jiné náhradní přechodové funkce, ovšem tato má celkem příjemné vlastnosti.

Při učení naší sítě tedy budeme postupovat tak, že si vezmeme náhodný vstup v_1, …, v_n, pro který známe správný výsledek $Y(v_1, \ldots, v_n)$. Postupně vypočítáme y_k a Y. Spočítáme rozdíl R našeho výsledku od správného výsledku

$$R = Y(s) - V(s)\,.$$

Pak vypočítáme nové hodnoty pro W_l, $w_{k,l}$ pomocí vztahů

$$W_l = W_l + \mu R\lambda Y \cdot (1 - Y)y_{2,l}$$

$$w_{k,l} = w_{k,l} + \mu R\lambda Y \cdot (1 - Y)W_l y_l \cdot (1 - y_l)v_k\,,$$

přičemž je μ rychlost, s jakou se síť učí. Pokud se síť učí rychle, může se stát, že přeskočí globální minimum a skončí v nějakém lokálním. Pokud se učí pomalu, tak učení bude trvat dlouho a opět se může stát, že skončíme v lokálním minimu.

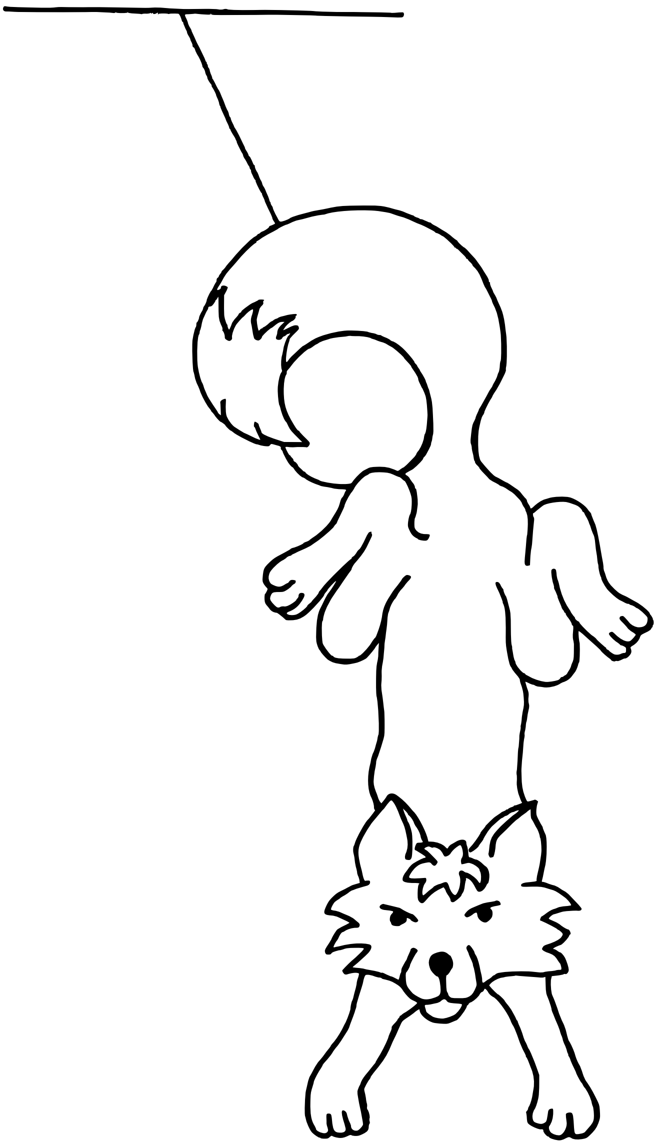
Předchozí učící postup stále opakujeme, dokud se chybová funkce neustálí. Pokud se ustálí na jiné hodnotě než nula, tak celý výpočet zahodíme a opět náhodně zvolíme w a p. A celý postup učení opakujeme.

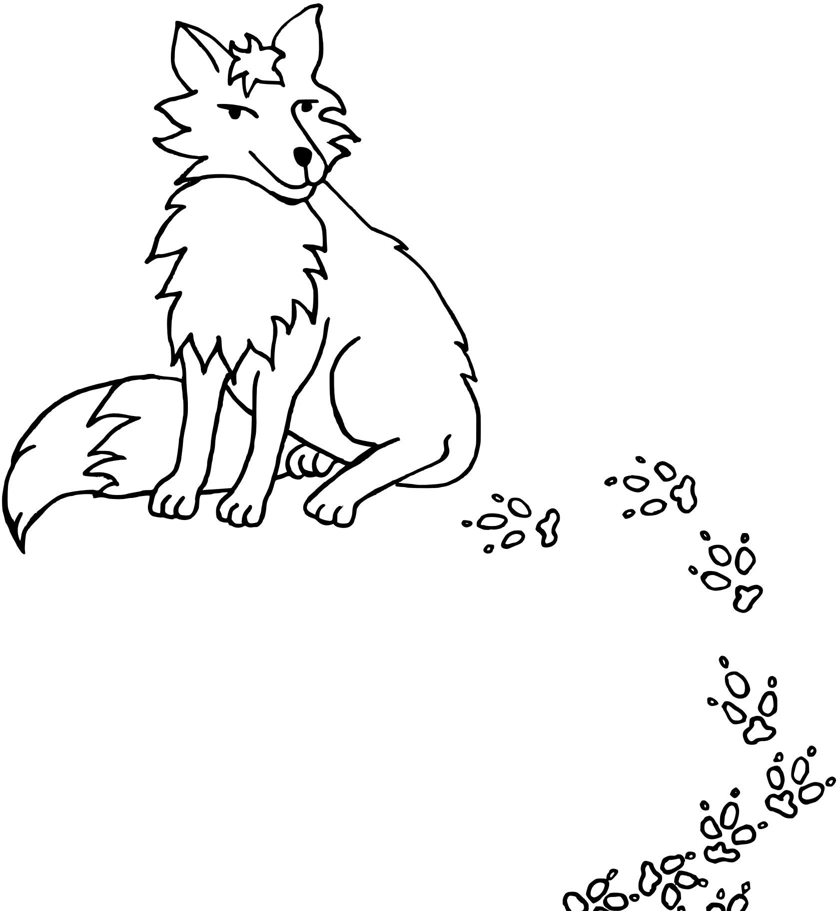
Výše popsaný způsob učení sítě se nazývá *back propagation*, metod, jak sítě učit, je však mnohem více. Neuronové sítě jsou v současné době hojně využívané a poznatky o nich jsou značně obsáhlé. Cílem tohoto článečku bylo vám je alespoň částečně představit. Pro úplné a poctivé představení bychom však potřebovali celou knihu. V případě jakýchkoli nejasností můžete vznášet dotazy či připomínky na mail `radim@matfyz.cz`.

(R)adim

Úloha 2.5 – Neuronová síť (4b)

Vytvořte program s neuronovou sítí, která bude mít dva vstupní neurony v_1 a v_2, skrytou vrstvu se třemi neurony a jeden výstupní neuron Y. Program v sobě bude mít učící smyčku, během které se síť naučí chovat podle funkce XOR, tedy pro $v_1 = v_2 = 0$ nebo $v_1 = v_2 = 1$ je $Y = 0$, pro $v_1 = 1$, $v_2 = 0$ nebo $v_1 = 0$, $v_2 = 1$ je $Y = 1$. Pro počáteční ladění programu můžete použít konstanty $\lambda = 10$, $\mu = 0{,}05$.





Adresa redakce:

M&M, OVVP, UK MFF
Ke Karlovu 3
121 16 Praha 2

Telefon: +420 221 911 235
E-mail: `mam@matfyz.cz`
WWW: `http://mam.mff.cuni.cz`

Časopis M&M je zastřešen Oddělením pro vnější vztahy a propagaci Univerzity Karlovy, Matematicko-fyzikální fakulty a vydáván za podpory středočeské pobočky Jednoty českých matematiků a fyziků.