

Pokroky matematiky, fyziky a astronomie

Otakar E. Kádner

Čtecí stroje

Pokroky matematiky, fyziky a astronomie, Vol. 4 (1959), No. 1, 74--79

Persistent URL: <http://dml.cz/dmlcz/137867>

Terms of use:

© Jednota českých matematiků a fyziků, 1959

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

ČTECÍ STROJE

OTAČAR E. KÁDNER

Státní ústav Centroprojekt, Praha

Automatisace evidenčních, statistických a výpočetních prací, prováděných elektronickými stroji, je brzděna mimo jiné i závažným problémem přípravy operace a vkládáním dat do stroje. Je známo, že tyto fáze pracovního cyklu elektronického počítače svým trváním a pracností daleko časově převyšují vlastní výpočet, trvající velmi krátkou dobu.

Čisté mzdy pro 50 000 zaměstnanců průmyslového podniku lze elektronickým agregátem vypočítat řádově za několik minut, avšak několik týdnů trvá příprava podkladů s děrováním do štítků nebo na papírový pásek.

Elektronický stroj vypočítá např. souřadnice 100 bodů za dvě minuty, ale děrování výchozích údajů a kontrola, která se dnes většinou musí provádět na kontrolním děrovači novým vyděrováním, trvá nejméně jednu hodinu.

A zde jsme u problému: snížit přípravu k výpočtům na elektronických počítačích tak, aby jejich využití bylo efektivní a pokud možno bez dlouhých prostojů.

Představa „mechanického oka“ zajímala mnoho výzkumných pracovníků již před desetiletími. Zde ovšem nejde o přenos obrazů jako při televizi, ale o vytvoření stroje, který by dovedl přečíst obyčejné písmo, „porozumět“ jeho obsahu a podle toho dále pracovat.

Stroj určený převážně pro administrativu by měl umět číst a zpracovávat obchodní doklady (např. třídit je) nebo knihovat obsažené údaje. Již v r. 1928 přihlásil rakouský profesor Tauschek podobný stroj k patentování. Dnešní stav elektronického zpracování informací umožnil, aby čtecí stroje vstoupily ze stadia vývoje do období praktického uskutečnění, takže je na místě podat přehled současného stavu na tomto poli, u nás dosud velmi málo známém.

Úkolem stroje na čtení by dnes především bylo přečíst údaje z originálního dokumentu, vyděrovat je pomocí děrovačích stroje do děrného štítku nebo na pás, popř. je přímo přenést do elektronického počítačového zařízení. Informace, kterou má stroj zpracovat, obsahuje někdy jen číslice, jindy však i písmena. Chceme ovšem, aby čtecí stroj pracoval rychle a absolutně bez chyby. Konstrukce stroje závisí na tom, jaké druhy písma má číst. Rozeznáváme v zásadě čtyři metody automatického čtení:

1. Normální tiskařské nebo strojové písmo,
2. normální tiskařské nebo strojové písmo s normovanými identifikačními znaky,
3. normované písmenné znaky a
4. rukopis.

Pokud nebude v dalším výkladu uvedeno jinak, vztahuje se pojem písmeno i na číslice uvedeného písma.

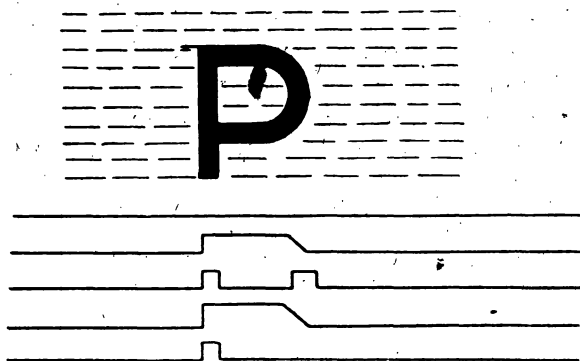
1. Na první pohled se zdá žádoucí, aby stroj mohl automaticky přečíst každý obyčejný druh písma. Co však je „obyčejné“ nebo „obyčejné“ písmo? Tiskařské i psací stroje používají sta různých druhů písem, které se od sebe podstatně liší tvarem i velikostí. Dosud je prakticky nemožné zkonstruovat stroj, který by mohl číst všechny druhy písem a je proto třeba se omezit na stroje pro „normální“ písmo, obsahující určitý počet vybraných druhů. Problém byl řešen srovnávací a impulsivní metodou.

Srovnávací metoda. Tato metoda spočívá v tom, že předloha se písmeno za písmenem srovnává se šablonou ve stroji, obsahující všechny typy druhů písem, které stroj může číst. Do neprůhledné šablony bývají znaky obvykle vyryty. Dokud se znak šablony přesně nekryje s předlohou, dopadá na fotonku rozptýlené světlo a stroj nereaguje. Jakmile nastá-

ne úplná shoda obou (obr. 1.), fotonka se zakryje, čímž se uvede v činnost registrační elektrický okruh. Vyobrazení č. 1 ukazuje jeden z prvních návrhů čtecího stroje podle Tauscheka [1]. Podobná zařízení navrhli *British Thompson-Houston* [2] a Handel [3], používající šablony ve tvaru desky. Goldberg [4] používal filmu. Bryce [5] má podobné zařízení, při němž se však písmeno předlohy opakuje na filmu v pěti různých nepatrně změněných polohách, aby se vyrovnaly případně tiskové a nastavovací nepřesnosti čteného písma. Koincidence registruje thyatron.

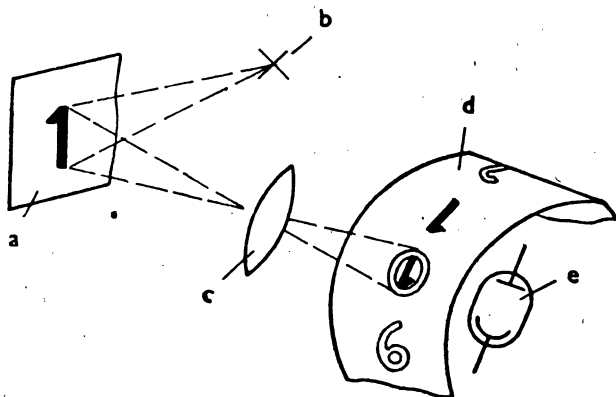
Srovnávací metoda automatického čtení je zdlouhavá, protože každé písmeno, které má být přečteno, se musí mechanicky srovnat s úplnou sadou písmen šablony nebo dokonce s několika sadami. Metoda není příliš spolehlivá, takže stroje tohoto typu se prakticky nerozšířily.

Impulsní kontaktní metoda. Při tomto způsobu se potišťená plocha rozloží na řadu velmi malých kontrolních plošek, z nichž každou pozoruje jedna fotonka, zapojená na mnohonásobný kontakt. Zakryje-li se jedna kontrolní ploška částí odečítaného písmene, zapojí se příslušný dílčí kontakt, kterým se připraví řada proudokruhů.



Obr. 2.

Metoda impulsních profilů (řad). Při tomto způsobu se každé písmeno ohmatává pohyblivým světelným bodem, takže příslušná fotonka vytvoří charakteristickou řadu nebo profil krátkých a dlouhých impulsů. Tato řada impulsů se porovnává s typickými profily, uloženými ve stroji. Souhlasí-li s některým, „rozpozná“ stroj příslušné písmeno a registruje je. Zworykin [7] používá pěti pohyblivých světelných bodů znázorněných kroužky



Obr. 1. Čtecí stroj dle Tauscheka. a – předloha, b – lampa, c – čočka, d – rotující buběn se šablonou, e – fotonka uvnitř bubny.

je zde jeden proudokruh pro každou možnou polohu písmene uvnitř pole. Fitch [6] používá 20 fotonek, z nichž každá je zapojena na jeden thyatron, jedno kontrolní relé a matici 12 × 12 neonových lamp. Během čtení je připraveno více proudokruhů, ale jen jediný se úplně propojí pro odečtené písmeno. Tato metoda je velmi rychlá, ale počet součástí stroje je značný. Je též pochybné, zda 20 fotonek postačí; každé zvýšení jejich počtu by však přineslo ještě větší zkomplikování stroje.

(1–5 v obr. 2). Každý světelný bod se pohybuje po čárkované linii; narazí-li na písmeno, vytvoří se v proudokruhu fotonky impuls. Řady nebo profily impulsů, vzniklé tímto způsobem, vidíme v druhé polovině obr. 2. Těchto pět řad definuje „ohmatané“ písmeno; nyní se zavedou do počítače, který zvláště registruje dlouhé a krátké impulsy. Odpočítané a zesílené impulsy procházejí funkční maticí a srovnávají se s typickými řadami, uloženými na povrchu magnetického bubnu (známá „paměť“ elektronických počítačích strojů). Původně bylo tohoto zařízení použito jako náhrady Braillova slepeckého písma, neboť odečtená písmena uváděla v činnost tlapkač a čtené věty byly tak převedeny na řeč. Zásadně by však tento stroj mohl tisknout nebo zpracovávat písmena v počítačím zařízení.

Flory [8] navrhl podobný stroj pro slepce, pracující s osmi světelnými body, ale také velmi komplikovaný. Shepard [9, 10] vytvořil pravděpodobně nejvíce vyvinutý stroj tohoto typu, používaný nyní již skutečně v praxi, např. ke čtení šeků a přenášení informací na děrné štítky, děrné nebo magnetofonové pásky. Pracovní rychlost je 3600 „slov“ za minutu, přesnost čtení je vyhovující za předpokladu, že písmena a číslice jsou dobře a čistě tištěny. Světelný bod vytváří během pohybu dokumentu strojem 30 impulsních profilů pro každý znak. Tyto profily se třídí speciálním elektronickým počítačím zařízením, které je identifikuje, ukládá do „paměti“ a dává dále na děrovací stroje apod. Nemůže-li tento značně složitý stroj nějaký znak indentifikovat (protože je např. znečištěn nebo nečitelný), dá signál „odmítnutí“.

Také Glaubermann [11] provádí čtení číslic impulsními profily. V jeho stroji se každý znak opticky zvětšuje a ohmatává řadou fotonek, zatím co jiné fotonky vyrovnávají obraz. Krátké a dlouhé impulsy každého profilu se sčítají a nakonec vyznačují dvoumístným identifikačním číslem, které prochází identifikační částí stroje s 35 magnetickými cívkami a maticí s 5×7 diodami. Nakonec se zapojí jeden z identifikačních proudokruhů pro číslice 0–9.

Na principu impulsních profilů byla vyvinuta celá řada strojů [12], avšak v rámci tohoto přehledného článku není možné je všechny popsat. Podotkneme jen, že anglický čtecí stroj systému SOLARTRON [12] má nyní přijít do sériové výroby. Všeobecně lze říci, že stroje pracující na principu impulsních profilů jsou ve vývoji nejpokročilejší a nejdokonalejší. Jsou však složité a poměrně drahé (řádově přes půl milionu Kčs).

2. Složitost a nákladnost impulsních strojů vedla k návrhu rozpoznávat písmena logicky uspořádanými identifikačními značkami. Pokud umístíme tyto značky mimo písmeno, není velikost a druh písma rozhodující, protože čtení se omezuje jen na rozpoznávání značek. Podle toho by mohlo být jediné čtecí zařízení universální a mohlo by číst písma všech velikostí, druhů a dokonce jazyků, ovšem za předpokladu, že se ve všech případech použije stejných identifikačních znaků. Takové zařízení by bylo mnohem jednodušší a lacinější než stroj impulsového typu. Ještě důležitější je snad ta okolnost, že takové zjednodušení zařízení by mohlo pracovat bez chyb.

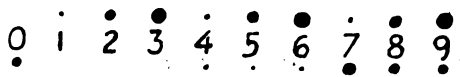
Nedostatkem metody je, že předlohy musí být vyhotoveny na speciálních sázečích nebo psacích strojích. Vnější identifikační znaky mohou příslušné písmeno charakterizovat svou velikostí, počtem nebo umístěním (posicí); toto poslední uspořádání lze strojem samostatně kontrolovat, čímž se vylučuje chyba.

Identifikační znaky různé velikosti. Heidinger [13] používá kruhové znaky pěti různých velikostí, jak názorně ukazuje obr. 3. Každé číslici od 0 do 9 je přiřazena jedna nebo dvě značky nad cifrou nebo pod ní. Dvě fotonky ohmatávají obě řady značek; každá z fotonek zapojuje pět relé laděných na příslušné intensity impulsů. Lze však pochybovat o tom, zda toto zařízení bude v praxi dost citlivé na poměrně malé velikostní rozdíly zna-

ků. Také samočinná kontrola není možná a náhodná skvrna (kaz v papíru) může způsobit chybu čtení.

Identifikační znaky podle počtu. Salchow [14] navrhuje podtrhnout každou číslici od 0 do 9 příslušným počtem čárek, které odečítá fotonka. Tohoto návrhu nelze zřejmě použít pro písmena, protože by bylo třeba velkého počtu čárek. Kontrola není též možná.

Identifikační znaky posiční. Hofgaard [15] používá k rozlišení písmen dvou skupin po pěti obdélníčcích. Šifrování je kontrolovatelné, protože horní a dolní řada obdélníčků se musí vzájemně doplňovat. Broido [16, 17, 18] používá šesti posic. Každá číslice od 0 do 9 je identifikována dvěma značkami. Leží-li značky těsně vedle sebe, spojují se v jedinou značku větší, aby bylo usnadněno jejich vyrytí. Mimo to je pod každou značkou startovní ryska, která má vyrovnat případné nesrovnalosti ve svislém uspořádání číslic. Devíti značkami lze šifrovat i písmena [20].



Obr. 3.

Psací stroj opatřený typy, upravenými podle tohoto návrhu, vytváří identifikační znaky pomocí speciální barevné pásky pouze na průklepu, zatím co originál ukazuje normální písmo.

3. Vytvoříme-li taková písmena, že každé má určité rozdílné znaky, může je vhodné automatické zařízení přečíst, i když nemají vnější identifikační značky podle odst. 2. Děje se tak opět na základě různých principů.

Systém různých ploch. Maul [21] navrhuje tisknout nulu slabě, jedničku silněji atd. Nula může např. překrývat 7% plochy, jednička 13% atd. po šesti procentech; devítka může zaujímat 61% potiskované plochy. Fotonka ohmatává potiskovaná místa a vysílá impulsy odstupňované síly, které ovládají příslušně laděné relé. Je otázka, zda tento způsob je dostatečně spolehlivý.

Systém různých frekvencí. France [22] používá číslic, které sestávají z vodorovných čárek, různě od sebe vzdálených. Fotonka vysílá následkem toho impulsy různé frekvence.

Systém svislých linií. Ducrocq [23] používá písmena, která se skládají se svislých a vodorovných čar. Obraz písmene se vrhá hranolem na dvě stínítka, z nichž jedno má vodorovné, a druhé svislé výřezy. Dvě fotonky jsou namířeny na tyto výřezy a registrují počet propouštěných čar.

Písmena s vlastními identifikačními znaky. Částí písmen lze účelně použít jako identifikačních znaků, které jsou pozorovány fotonkami. Tak např. Maul [21, 24] používá tři až devíti kontrolních bodů pro číslice a jedenácti pro písmena. Pro každý kontrolní bod je umístěna jedna fotonka. Bryce [25] používá čtyř kontrolních bodů pro deset číslic. Systém *Standard telephones* [26] má pět kontrolních bodů, které se postupně osvětlují. Vroom [27] používá šesti kontrolních bodů pro číslice; body jsou opticky zvětšovány a pozorovány šesti fotonkami. Roet [28] rozděluje pole na 15 malých čtverců, číslice i písmena uspořádává tak, že každé pokrývá určitým počtem čtverečků (obr. 4). Fotonky sledují pouze deset čtverečků (střední a pravý sloupec).



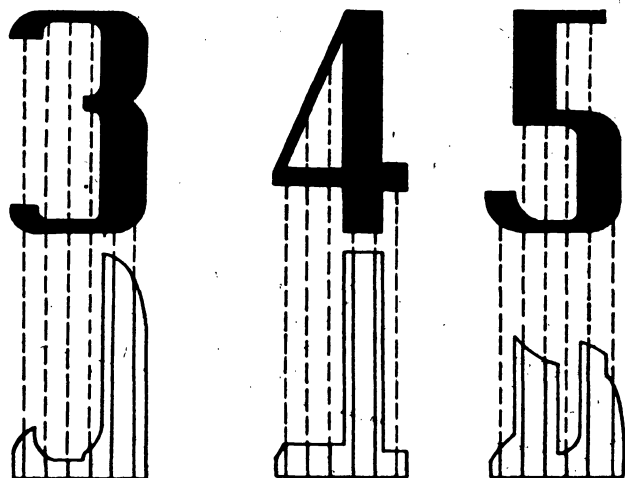
Obr. 4.

Paris [29] rozděluje pole na deset vodorovných kontrolních plošek, ale jen úzký svislý sloupec sledují fotonky. Nejhořejší značka je startovní. Podobné zařízení navrhl Crosfield [30]. Ayres [31] uspořádal čtyři kontrolní body uvnitř obrysu číslic, které se odečítají rotující deskou s otvory. Zařízení

je citlivé, avšak pracuje pomalu. Všechny dosud popsané metody nelze při provozu samočinně kontrolovat.

Dickinson [32] používá jediného většího kontrolního bodu v různé výšce. Zařízení je kontrolovatelné samočinně, avšak lze pochybovat o jeho spolehlivosti, neboť kontrolní bod nemůže být větší než jedna desetina výšky číslice. Broido rozděluje tiskové pole na deset vodorovných kontrolních ploch.

Systém magneticky tištěného písma. Největší potíž spolehlivého automatického čtení je v poruchách, které vznikají nečistotou, špatným tiskem nebo kazy papíru. Četní badatelé se proto snažili vyvinout takovou metodu, na kterou by tyto vady neměly vliv.



Obr. 5.

K tiskařské barvě lze např. přimísit mikroskopicky malé železné částičky, které se po tisku zmagnetisují. Takto tištěné znaky možno pak magnetickým ohledáním identifikovat [12]. Magnetické značky nejsou ovlivněny skvrnami, kazy papíru, přerazítkováním atd. Za jistých okolností mohou být dokonce „čteny“ i pod nálepkami nebo poštovními známkami. Mohou je však rušit kovové částice které jsou často v papíře obsaženy.

Čtecí stroj vyvinutý v *Stanford Research Institute* [34, 35] spočívá na speciálně formovaných, magneticky tištěných číslicích (obr. 5). Při pohybu předlohy zleva doprava sčítá zařízení celkový magnetický impuls a vytváří elektrický signál, znázorněný profilem v dolní polovině obr. 5. Tyto signály 200 mV se zesílují na 50 V a přivádějí na identifikační proudokruhy; jeden z nich se po srovnání nabije kladně, ostatní zůstanou negativní.

Jsou-li dva okruhy současně kladné, vyšle stroj odmítavý signál, protože jde o chybu. Tohoto stroje se již v praxi používá pro třídění šeků. Ve srovnání se světelnou metodou má magnetické čtení velké výhody, ale i určité vady; potřebuje např. speciální tiskařské barvy nebo kopirovací papíry. Elektrické zapojení je značně složité, avšak mohlo by být zjednodušeno použitím normovaných písmen.

4. Teoreticky je možné rozšířit impulsní metodu i na čtení rukopisu, ale prakticky je to příliš obtížné. Proto takový stroj dosud není znám. Bylo však navrženo [36] psát rukopisné záznamy na dřevný štítek tak, aby každá číslice byla identifikována podle své polohy na štítku. Pro automatické čtení pracovních záznamů a výkazů apod. lze rozpisovat

číselné údaje tak, aby padly na různá předem stanovená místa. Např. číslo 5637 lze napsat ve tvaru

3

5 .

6 .

.7

a fotonka registruje číslice podle police.

Souhrnem můžeme říci, že čtecí stroje jsou v nynější době teoreticky připraveny a existuje několik typů, které jsou v praktickém provozu. Stroje založené na čtení určitého druhu písma nebo písma označeného identifikačními značkami, jsou poměrně jednoduché a levné konstrukce. Universální stroje, pracující většinou na principu impulsních profilů, jsou značně složitá a drahá.

Celkem můžeme říci, že doba, kdy lidské oko a ruka budou vyřazeny z přípravy operací, která zabírá neúnosně velkou část celkové doby zpracování informací, je velmi blízká.

Literatura a patentní spisy

- [1] Tadschek, rakouský patent 116799, něm. patent 662417, USA-patent 2026329.
- [2] *British Thompson-Houston*, angl. patent 396504.
- [3] Handel, GEC, patent USA 1 915 993.
- [4] Goldberg, Zeiss, angl. patent 288580, USA pat. 1 838 389, franc. patent 657787.
- [5] Bryce, IBM, angl. patent 504719.
- [6] Fitch, IBM, USA-par. 2682043.
- [7] Zworykin, RCA, USA-pat. 2616983.
- [8] Flory, RCA, USA-pat. 2615992.
- [9] Shepard, IMC, USA-pat. 2663758.
- [10] D. H. Shepard — C. C. Heasley, *Electronics*, N. Y. 1955.
- [11] M. H. Galubermann, *Electronics*, N. Y. 1956.
- [12] "*Electronics and Banks*" Peat, Marwick, Mitchell Co, N. Y. 1956.
- [12A] *SOLARTRON Electronic Group Ltd.*, Thames Ditton, Surrey, Anglie, "*SOLARTRON Reading Machines*".
- [13] Heidinger, IBM, USA-pat. 2362004.
- [14] Salchow, něm. patent 671484.
- [15] Hofgaard, USA-pat. 2337533.
- [16] Broido, angl. patent 538016.
- [17] Broido, angl. patent P. A. 4512/54.
- [18] Broido, angl. P. A. 32533/54.
- [19] Broido, BTM, angl. P. A. 3445/56.
- [20] Broido, BTM, angl. P. A. 32534/54.
- [21] Maul, angl. patent 288327, 528849; něm. patent 688181, francouzský patent 864034.
- [22] France, angl. patent 639054.
- [23] Ducrocq, franc. patent 1104482.
- [24] Maul, IBM, franc. patent 863916, 864034, 866394; angl. patent 528849, 535031; USA-pat. 2285296, 2294679.
- [25] Bryce, angl. patent 360044; něm. patent 542026; USA-pat. 2063481.
- [26] *Standard Telephones*, angl. patent 347134.
- [27] Vroom, *Bell Telephone*, USA-pat. 2723308.
- [28] Roet, angl. patent 376283.
- [29] Paris, IBM, angl. patent 538604.
- [30] Crosfield, angl. patent 650536.
- [31] Ayres, IBM, USA-pat. 2131911.
- [32] Dickinson, IBM, angl. patent 522956, 522992, USA-pat. 2261542.
- [33] *ERMA, Electronic Recording Machine*, Accounting, *Stanford Research Institute*, Menlo Park, Cal. USA.
- [34] Eldredge, Kampfhofner a Wendt, *Automatic Input of Business Date Processing Systems*. Stanford Res. Inst. Hjen 1956.
- [35] Bull Co., franc. patent 104874, 1002613.