

JAZYK GDL NA SPRACOVANIE VEDECKÝCH DÁT

ŠECHNÝ, Martin (SK)

Abstrakt. GNU Data Language (GDL) je jazyk na spracovanie vedeckých dát a zároveň prostredie na spúšťanie programov v tomto jazyku. GDL je slobodný softvér kompatibilný s komerčne licencovaným Interactive Data Language (IDL). GDL je platformovo nezávislé prostredie a využíva iné dostupné nainštalované knižnice a aplikácie. Jazyk GDL umožňuje spracovávať vstupy z klávesnice, dátové súbory, obrázky a dokáže vizualizovať dáta tabuľkami, grafmi, obrázkami. GDL je efektívny pri numerickej analýze dát, vektorovej reprezentácii, použití matematických funkcií a procedúr. Tento nástroj je vhodný pre široké použitie vo vede, výskume, aj ako alternatíva k známym matematickým a vizualizačným nástrojom.

Kľúčové slová. GDL, IDL, vedecké dáta, programovanie, vizualizácia.

GDL LANGUAGE FOR SCIENTIFIC DATA PROCESSING

Abstract. GNU Data Language (GDL) is a language for scientific data processing and also the environment for launching programs in that language. GDL is a free software that is compatible with commercially licensed Interactive Data Language (IDL). GDL is a platform-independent environment and uses other available libraries and applications installed. GDL language enables to process keyboard input, data files, images and can visualize data tables, charts, pictures. GDL is effective in the analysis of numerical data, vector representation, the use of mathematical functions and procedures. This tool is suitable for wide use in science, research, and as an alternative to known mathematical and visualization tools.

Key words and phrases. GDL, IDL, scientific data, programming, visualization.

1 Úvod

GNU Data Language (GDL)¹ je jazyk na spracovanie vedeckých dát a zároveň prostredie (interpreter a inkrementálny prekladač) na spúšťanie programov v tomto jazyku. GDL je slobodný softvér kompatibilný s komerčne licencovaným Interactive Data Language (IDL)². Historický vývoj jazyka začal v roku 1977³.

GDL je platformovo nezávislé prostredie, programované pre systémy spĺňajúce štandard POSIX. Využíva dostupné nainštalované knižnice a aplikácie pre rozšírenie svojich funkcií a možností, alebo ako lepšie optimalizované implementácie štandardných, ale menej efektívnych funkcií. Inštalácia GDL je vďaka tomu modulárna, prispôsobiteľná použitiu. GDL sa stále vyvíja a prispôsobuje sa jazyku IDL v najnovšej verzii. GDL je programované v C++ s použitím viacerých ďalších programovacích a matematických jazykov. [1], [2]

Jazyk GDL umožňuje spracovávať vstupy z klávesnice, dátové súbory, obrázkové súbory, a to jednotlivo, aj hromadne. Výstup môže byť textový v terminálovom

1 GDL 0.9.1 (29.3.2011) <http://gnudatalanguage.sourceforge.net>

2 ITT Visual Information Solutions IDL 8.0 (30.9.2010)
<http://www.itvis.com/language/en-us/productservices/idl.asp>

3 Wikipedia: IDL: [http://en.wikipedia.org/wiki/IDL_\(programming_language\)](http://en.wikipedia.org/wiki/IDL_(programming_language))

okne, grafický, alebo do súboru. Je možné vizualizovať dáta tabuľkami, grafmi, obrázkami. GDL má široké možnosti matematického spracovania dát – numerická analýza, vektorová reprezentácia, výpočet funkcií a procedúr.

GDL, ako jazyk a zároveň prostredie, je vhodný pre široké použitie vo vede, výskume, aj ako alternatíva k známym matematickým a vizualizačným nástrojom.

2 Inštalácia

Prostredie GDL je programované multiplatformovo, pre systémy spĺňajúce štandard POSIX (napr. UNIX, BSD, GNU/Linux, Mac OS X), po doplnení potrebných súčastí môže fungovať aj v MS Windows.

Sú k dispozícii spustiteľné balíky⁴ pre viaceré distribúcie GNU/Linux, BSD, Mac OS X. Základná inštalácia hotového balíka je jednoduchá a rýchla. Prostredie GDL hneď môžeme vyskúšať v interaktívnom príkazovom riadku.

Ak nevyhovuje ani jeden spustiteľný inštalačný balík, musíme nainštalovať GDL zo zdrojového kódu⁵ po preklade. Pri inštalácii musíme vyhovieť závislostiam. GDL nutne potrebuje knižnice *plplot*, *gsl*, *readline*. Podľa účelu použitia GDL si vyberieme voliteľné knižnice *Xlib*, *Magick++*, *FTW*, *netCDF*, *HDF*, *GRIB API*, *UDUNITS*, *wxWidgets*, *libproject*, *GSHHS*, *Python*, *CMSVLIB*, *MPFIT*, *Astron lib*, *TeXtoIDL* a iné. Na preklad zdrojového kódu GDL a ručnej inštalácii je potrebné mať dobrú znalosť operačného systému a programovacích jazykov, lebo je potrebné zvoliť správne poradie prekladu a inštalácie knižníc a tie v správnych verziách. Hlavnou nutnou knižnicou je *gsl* (*GNU scientific library*), do ktorej sa pripájajú voliteľné knižnice a spolu s nimi poskytujú funkcie a procedúry volané prekladačom/interpreterom prostredia GDL. Často je potrebné ručne nastavovať viaceré parametre prekladu vzhľadom na použitý operačný systém. Medzi podporné knižnice patria knižnice pre algebru (*blas*, *cblas*, *lapack*) prekladané ako súčasť *gsl*, alebo pripojené zvlášť.

Pri väčšom počte knižníc trvá preklad a inštalácia bežne niekoľko hodín a spolu s naštudovaním dokumentácie ubehne aj niekoľko dní. Inštalácia obsahuje viaceré pripravené testy funkčnosti a optimalizácie prostredia, knižníc, funkcií.

3 Spustenie prostredia GDL

Prostredie GDL sa môže spustiť v terminálovom okne (príkazovom riadku):

```
$ gdl
```

Nainštalované môže byť aj grafické prostredie ako samostatná nadstavba *gdl*, alebo ako modifikovaná aplikácia *gdl*. Dokumentácia ku GDL nie je dostatočne podrobná alebo zrozumiteľná, preto ďalej uvádzam postup pri prvom spustení.

4 <http://gnudatalanguage.sourceforge.net> (download)

5 <http://sourceforge.net/projects/gnudatalanguage/>

Prostredie GDL potrebuje mať nastavené systémové premenné. Existujú však odlišné inštalácie GDL/IDL, preto sa odporúča nastaviť duplicitné premenné s rôznymi názvami, aby sa vyhovelo rôznym situáciám, napr.:

```
~/bashrc (~/.profile):
```

```
export GDL_DIR=/opt/gdl
export IDL_DIR=/opt/gdl
export GDL_PATH=/opt/gdl/pro
export IDL_PATH=/opt/gdl/pro
export GDL_STARTUP=~/.gdl/gdlrc
export IDL_STARTUP=~/.gdl/idl_startup/pro
```

GDL/IDL hľadá súbor s nastaveniami, ktorý takisto môže mať rôzne názvy:

```
~/gdl/gdlrc
~/gdl/idl_startup.pro -> gdlrc
```

V tomto súbore môžu byť direktívy pre beh prostredia GDL, príkazy operačného systému, príkazy jazyka GDL – vykonávajú sa pri spustení prostredia. Dôležité je vložiť do premennej *PATH* cestu k funkciám, procedúram, programom, ktoré má prostredie GDL vedieť nájsť podľa mena:

```
!PATH='/opt/gdl/' ; príklad cesty ku knižnici GDL
!PATH=expand_path('+'+!PATH) ; rozvinúť cestu do stromu
```

Pri spustení GDL môžeme zmeniť aktuálny priečinok príkazom *cd*, alebo spustiť GDL na takom mieste, kde máme svoje súbory.

Nainštalované prostredie GDL obsahuje zhruba 100 súborov *.pro* so základnými funkciami/procedúrami jazyka. Na toto miesto môžeme nakopírovať ďalšie súbory, alebo pripojiť svoju cestu do premennej *PATH*.

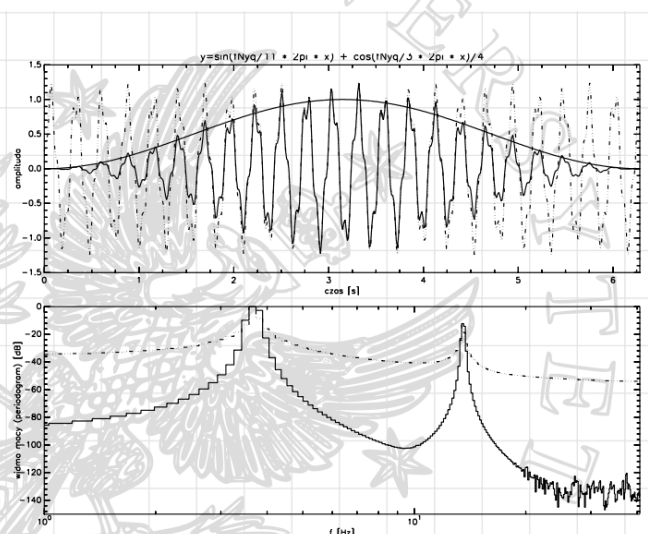
Príkladom je knižnica *Astron lib*⁶ pre astronómov, fyzikov. Podobných knižníc je mnoho. Jazyk GDL ešte nie je tak známy ako starší jazyk IDL, preto sa ľahšie hľadajú knižnice s kľúčovým slovom *IDL*. Autor programu v jazyku GDL/IDL sa môže rozhodnúť, aké licenčné podmienky požaduje pre svoje zverejnené programy a je výhodné, ak taký program funguje v ľubovoľnom prostredí, teda aj v GDL.

6 Goddard Space Flight Center IDL Astronomy User's Library (Astron lib)
<http://idlastro.gsfc.nasa.gov>

```

hann_demo.pro
1 pro periodogram, czas, sgnl, _extra = ex
2 n = n_elements(czas)
3 dt = czas[1]-czas[0]
4 f = findgen(n/2+1) / n / dt
5 amp2 = abs((fft(sgnl)) / (n/2))2
6 amp2[1:n/2-1] *= 2 ; uwzgl. częstotści ujemnych
7 moc_db = 10*log10(amp2/max(amp2))
8 ; norm. do max więc bez norm. do sumy wag w oknie
9 oplot, f, moc_db, psym=10, _extra = ex
10 end
11
12 pro hann_demo
13 n = 512 dt = 1./(n-1) * 2 * PI
14 nyquist = .5 / dt
15
16 czas = findgen(n) * dt
17 sgnl = sin(nyquist/11 * 2*PI * czas) +
18 cos(nyquist/3 * 2*PI * czas) / 4
19 hann = hanning(n)
20
21 IP.MULTI = [0,1,2] & IX.STYLE = 1
22
23 plot, linestyle=4, czas, sgnl,
24 xtitle="czas [s]", ytitle="amplituda",
25 title="wynik sin(nyq/11 * 2pi * x) +
26 cos(nyq/3 * 2pi * x)/4"
27 oplot, linestyle=0, czas, sgnl+hann
28 oplot, czas, hann
29
30 plot, [0, nyquist], [0, -150], /nodata,
31 xtitle="f [Hz]", /xlog,
32 ytitle="widmo mocy (periodogram) [dB]"
33 periodogram, linestyle=4, czas, sgnl
34 periodogram, linestyle=0, czas, sgnl * hann
35 end

```



Obrázok 1: Ukážka programu v jazyku GDL s grafickým výstupom⁷

Obrázok 2: GDL funguje v MS Windows + Cygwin⁸

7 (upravený) <http://gnudatalanguage.sourceforge.net/screenshots/screenshot06.png>

8 (upravený) <http://gnudatalanguage.sourceforge.net/screenshots/screenshot09.png>

4 Jazyk GDL

Program v GDL sa píše ľahko. Jazyk GDL sa podobá čiastočne jazyku FORTRAN, čiastočne jazyku C, využíva sa množstvo výhod podporných knižníc prostredia, hlavne matematické výrazy. V príkazovom riadku je dostupná pomoc, v grafickom prostredí aj základná príručka, príklady, zoznamy príkazov. Stručný návod pre začiatočníkov nájdeme na internete⁹, ako aj obsiahly zoznam zdrojov pre široké použitie jazykov GDL/IDL¹⁰. Funkcie, procedúry, programy a dokumentácia časom pribúdajú, lebo ich tvorí odborná verejnosť pri svojej práci a obvykle sú sprístupnené všetkým. Podrobná dokumentácia IDL pomôže aj pri práci s GDL. [3], [4], [5]

Nasleduje stručne ukázaná syntax a zopár základných príkazov z jazyka GDL:

Nerozlišujú sa malé a veľké písmená. Príkaz sa ukončuje koncom riadka (nie bodkočiarkou). Príkaz operačného systému začína znakom \$. Systémová premená sa použije s uvádzajúcim znakom !. Komentár začína na riadku znakom ;.

```
<príkaz>, <parameter> ; komentár
```

Dátové typy sú automatické, stačí premennú prvý raz použiť, nie je nutné ju deklarovať. Celé číslo (int) má štandardne 2 B, alebo veľkosť podľa operačného systému. Niektoré parametre funkcií alebo procedúr začínajú znakom /.

```
A=5 ; priradenie hodnoty do premennej A
print, A ; výpis premennej
help, A ; výpis informácie o premennej
B=[1,2,3] ; vektor/pole/matica
print, B[0:1] ; vypíš prvky poľa B s indexom 0 až 1
C=A+5 ; matematická operácia
S='prvy'+ " druhý" ; operácia s textovým reťazcom
openr,subor,'in.txt',/get_lun ; otvorenie súboru na čítanie
readf,subor,S ; načítanie súboru do premennej S
print,subor ; výpis súboru
free_lun,subor ; zavrieť súbor
```

Vektorové/maticové operácie sú vykonávané automaticky, ak je to potrebné, upravia sa rozmery (dimenzie aj počet prvkov) poľa a jeho dátový typ.

Programový súbor má príponu *.pro*, končí sa príkazom *end*. Program sa musí najprv preložiť a potom sa môže vykonať:

```
.compile <program> ; preklad
.run <program> ;(preklad a) vykonanie programu
<program> ; vykonanie programu (procedúry)
```

Blok je ohraničný kľúčovými slovami *begin*, *end*. Podprogram môže byť vo forme procedúry alebo funkcie:

⁹ Sven Geier, GDL intro, 2006: <http://www.sgeier.net/tools/GDL-intro.pdf>

¹⁰ IDL resources: <http://www.astro.virginia.edu/class/oconnell/astr511/IDLresources/astro-dept-idlresources.html>

```

pro program, x                ; procedúra
    x=x+1
    print,x
    return

end

function funkcia, x          ; funkcia
    x=x+1
    return,x

end

```

Jazyk umožňuje používať podmienku, vetvenie, skok, cyklus, napr.:

```

if A eq B then print,S
if A gt B then begin
    print,A
    print,B
endif else begin
    print,S
endelse
for I=1,10,1 do print,I
while A eq B do A=A+1
repeat A=A+1 until A eq B

```

Spracovanie parametrov programu:

```

if(n_params(0) eq 0) then print,'Neboli zadane parametre'

```

GDL je veľmi vhodný na spracovanie obrázkov, čo sa využíva hlavne v astronómii, ale napr. aj v medicíne:

```

read_jpeg, '1.jpg', image, true=1      ; načítanie obrázka
ysize=size(image,/dimensions)        ; zistenie rozmerov
window, 0, xsize=ysize[1], ysize=ysize[2], title=Obrazok ; okno
tvsc1, image, true=1                  ; zobrazenie obrázka
obrazok=fltarr(3, 300, 300)           ; deklarácia obrázka v reálnom
                                        poli (RGB, x, y)

```

Knižnica obsahuje veľa funkcií/procedúr pre transformáciu obrázkov, podobne ako kvalitné grafické aplikácie. Tieto operácie sa dajú algoritmizovať a sú vykonávané efektívne.

Ďalej existujú knižnice pre prácu s rôznymi formátmi súborov, s databázou, na paralelné vykonávanie programov, simulácie. Dáta možno vypísať do terminálu na obrazovke, zapísať do súboru, vyrobiť z nich tabuľku, alebo zobraziť ako obrázok pomocou rôznych vizualizačných nástrojov, napr. Plplot.

5 Aplikácie jazyka GDL

Veľká skupina používateľov jazyka GDL/IDL pracuje vo fyzike, špeciálne v astronómii. Okrem knižnice *Astron lib*¹¹ je možné použiť GDL ako API k špecializovaným lokálnym alebo webovým aplikáciám, napr. na ovládanie astronomického programu Aladin¹², prístup k virtuálnym observatóriám¹³. Jazyk GDL/IDL sa používa na spracovanie a analýzu obrázkov z astronomických pozorovaní, napr. snímky Slnka, vzdialených vesmírnych objektov, nočnej oblohy.

GRID je distribuovaný výpočtový systém, pozostávajúci zo samostatných počítačov, alebo klastrov a navonok sa tvári ako virtuálny superpočítač, ktorý môže poskytnúť používateľovi ohromné výpočtové zdroje. GDL je univerzálne prostredie, dá sa nainštalovať aj na slabší jedno-procesorový počítač, aj na virtuálny počítač. GDL v GRIDE vie obslužiť silne paralelizovateľné úlohy a vykonať veľmi efektívne vektorové/maticové operácie definované v matematických knižniciach. Tieto možnosti robia z GDL výkonný simulačný nástroj. [6], [7]

Jazyk GDL sa používa vo viacerých vedných oblastiach aj napr. na spektrálnu analýzu, fotometriu, redukovanie obrázkov, fitovanie.

6 Záver

GNU Data Language (GDL) je jazyk na spracovanie vedeckých dát. Je vhodným slobodným softvérovým nástrojom pre vedu a výskum, kompatibilný s komerčne licencovaným Interactive Data Language (IDL).

GDL možno porovnať s matematickými nástrojmi¹⁴, napr. *Octave*, *Scilab*, *Pylab*, alebo komerčne licencovanými *Matlab*, *Mathematica*, ďalej so simulačnými nástrojmi, napr. *GENESIS*. Veľký prínos má GDL s knižnicou *Astron lib* pre fyzikov, astronómov, kde môžu analyzovať a spracovávať obrázky vesmírnych objektov, čo sa dá porovnať napr. s *AIPS*, *IRAF*.¹⁵

Pri použití GDL v GRIDE máme v rukách nástroj so silným výpočtovým výkonom a poradíme si aj veľmi veľkým objemom vstupných dát.

11 Goddard Space Flight Center IDL Astronomy User's Library (Astron lib)
<http://idlastro.gsfc.nasa.gov>

12 Aladin <http://aladin.u-strasbg.fr/aladin.gml>

13 Virtual Observatory lib <http://www.ctio.noao.edu/~chrism/VOLib>

14 Matematický softvér

<http://www.personal.kent.edu/~rmuhamma/Mathematics/mathsoftware.html>

15 Astronomický softvér <http://tdc-www.harvard.edu/astro.software.html>

Literatúra

[1] Arabas, .S: GDL - GNU Data Language, Free and Open Source Developers' European Meeting 2011 <http://www.igf.fuw.edu.pl/~slayoo/gdl-fosdem-2011-talk.pdf>

[2] COULAIS, A. - MERMET, T.: Status of GDL - GNU Data Language, Astronomical Data Analysis Software and Systems XIX, ASP Conference Series, Vol. XXX, 2009 arXiv:1101.0679v1 [astro-ph.IM] 4 Jan 2011 <http://arxiv.org/abs/1101.0679>

[3] ITT Visual Information Solutions: Getting Started with IDL, 2007 1107IDL70GS <http://www.astro.virginia.edu/class/oconnell/astr511/IDLresources/getting-started-IDL-v7.0.pdf>

[4] ITT Visual Information Solutions: IDL Quick Reference, 2009 0509IDL71QR http://www2.fz-juelich.de/vislab/software/idl_docs/quickref.pdf

[5] Research Systems: IDL HandyGuide, 2000 <http://www.astro.virginia.edu/class/oconnell/astr511/IDLresources/IDL-handiguide-v5.4.pdf>

[6] WANG, M., FOX, G., PERCE, M.: Grid-based Collaboration in Interactive Data Language Applications, In Proceedings of ITCC (1)'2005, 335-341 2005 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.7212&rep=rep1&type=pdf>

[7] SHASHARINA, S. - VEITZER, S.: GRIDL: High-Performance and Distributed Interactive Data Language, <http://www.txcorp.com/pdf/publications/GRIDL.pdf>

Kontaktná adresa

Martin ŠECHNÝ (Mgr.), Stredná priemyselná škola elektrotechnická, Plzenská 1, 080 47 Prešov, martin.sechny@shenk.sk

