

Obliczenia Naukowe

Wykład 11: Pakiety do obliczeń:
naukowych i inżynierskich
Przegląd i porównanie

Bartek Wilczyński

29.5.2017

Plan na dziś

- Pakiety do obliczeń: przegląd zastosowań
- różnice w zapotrzebowaniu: naukowcy, inżynierowie, statystycy/medycy
- Matlab/octave/scipy – obliczenia numeryczne
- Mathematica/Maxima/Sympy – obliczenia symboliczne
- Sage i SageCloud – zintegrowany pakiet opensource
- Excel?

Typowi użytkownicy pakietów obliczeniowych

- Inżynierowie i projektanci (budownictwo, lotnictwo, motoryzacja, itp.)
- Naukowcy doświadczalni (fizycy, chemicy, materiałoznawcy, itp.)
- Statystycy (zastosowania w medycynie, ekonomii, biologii molekularnej, psychologii, socjologii, ubezpieczeniach, itp.)
- Matematycy (przede wszystkim matematyka stosowana)

Obliczenia naukowe

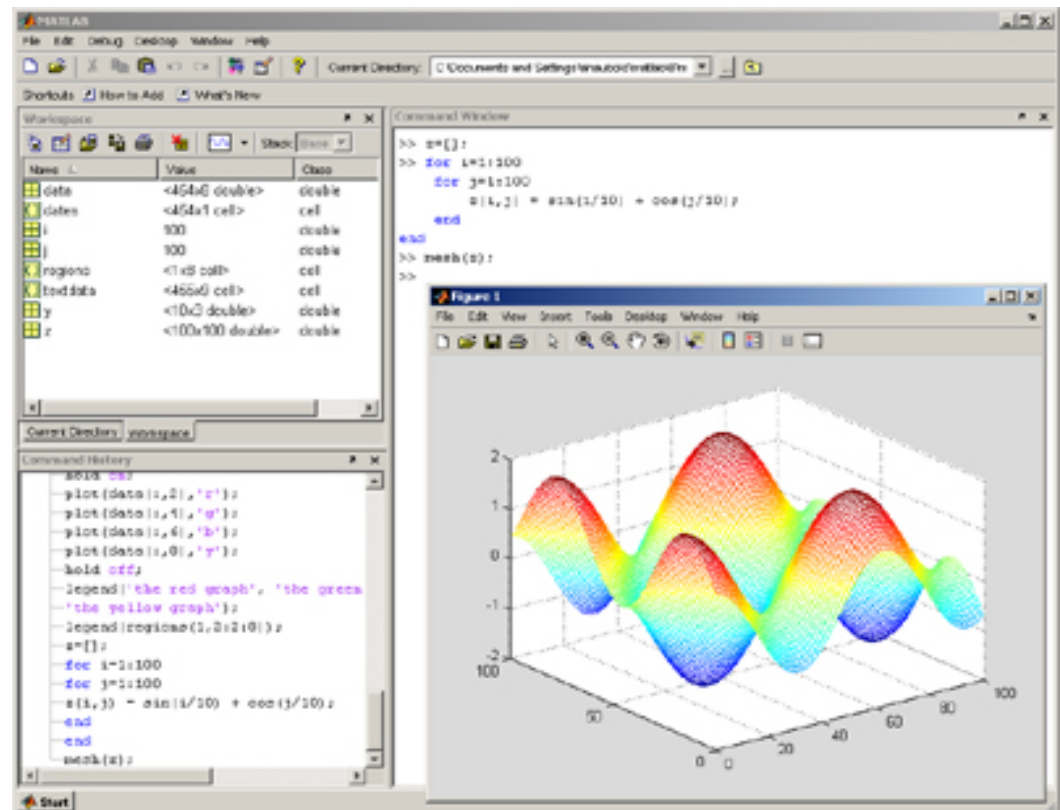
- Komputer jako “potężniejszy kalkulator”
- W zasadzie wszystko można zaprogramować samemu, ale każdemu mogą się przydać:
 - Interfejs użytkownika łatwiejszy niż typowego kompilatora
 - Możliwość zaawansowanej grafiki
 - Dobrze przetestowane standardowe procedury
 - Interfejsy do urządzeń
 - Wsparcie fachowców

Matlab i pakiety “inżynierskie”

- Rozwijany w latach 70'tych przez Cleve Moler'a jako narzędzie dla studentów informatyki, aby nie musieli używać zaawansowanych bibliotek fortranu
- Firma mathworks powstaje w 1984 i wydaje pierwszą wersję Matlab'a
- Najpopularniejszy wśród inżynierów, dobre całki numeryczne, rozwiązywanie równań i wykresy (również 3d)
- Bardzo popularny także do przetwarzania sygnałów i symulacji (simulink)
- Licencja komercyjna – niedrogi dla studentów, droższy dla uczelni, bardzo drogi dla przemysłu

Toolbox'y Matlab'a

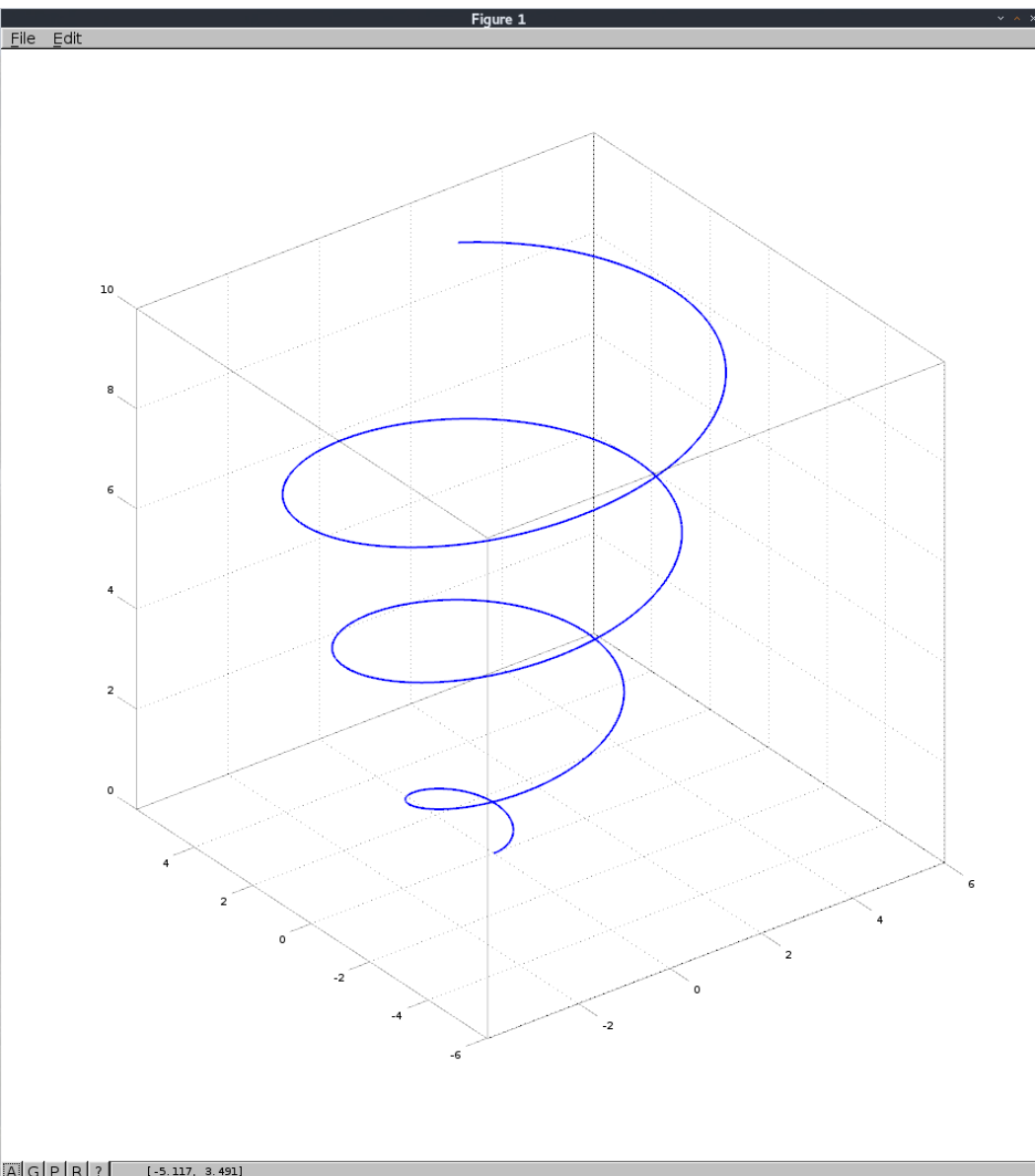
- Wiele dodatkowych (płatnych) bibliotek dla specjalistów
 - Symbolic math
 - Image processing
 - Financial toolbox
 - Bioinformatics
 - Optimization
 - SimBiology



Alternatywy openSource

- GNU Octave (rozpoczęty w 1988, wydania od 1992, rozwijany przez John'a W. Eatona, chemika z University of Wisconsin-Madison)
 - W zasadzie kompatybilny z Matlab'em
 - John W. Eaton Inc. - consulting
- Scipy stack – zestaw bibliotek python'a do obliczeń naukowych
 - Wiele bibliotek, rozwijanych przez niezależne grupy
 - System pakietów, edytor i dystrybucja organizowana przez firmę Enthought, również komercyjne dystrybucje i consulting
 - Wiele konferencji tematycznych dla naukowców i pracowników przemysłu - także źródło dochodu

Interfejs Octave



```
octave

~ » octave
GNU Octave, version 3.8.1
Copyright (C) 2014 John W. Eaton and others.
This is free software; see the source code for copying conditions.
There is ABSOLUTELY NO WARRANTY; not even for MERCHANTABILITY or
FITNESS FOR A PARTICULAR PURPOSE. For details, type 'warranty'.

Octave was configured for "x86_64-unknown-linux-gnu".

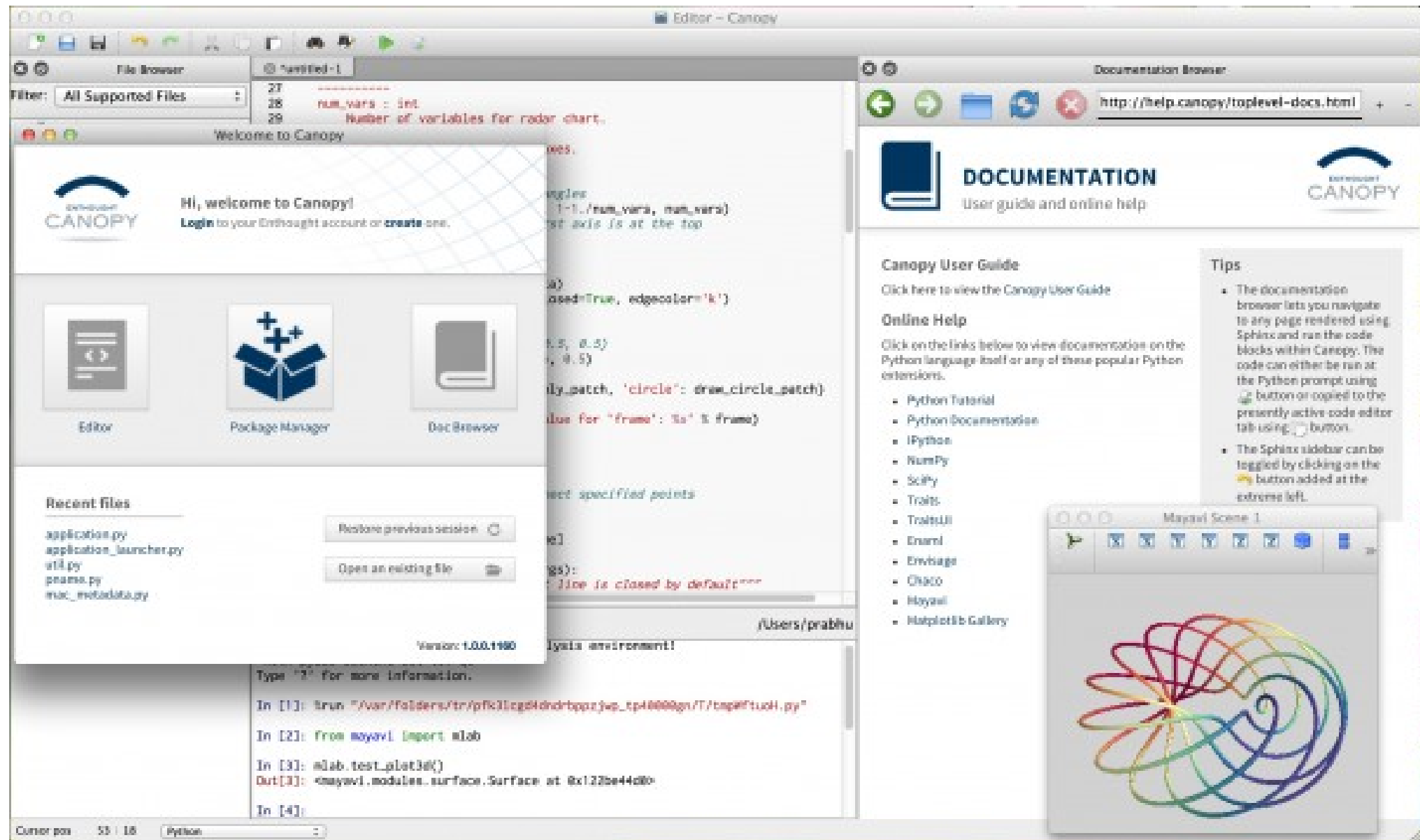
Additional information about Octave is available at http://www.octave.org.

Please contribute if you find this software useful.
For more information, visit http://www.octave.org/get-involved.html

Read http://www.octave.org/bugs.html to learn how to submit bug reports.
For information about changes from previous versions, type 'news'.

octave:1> t=[0:0.01:20];
octave:2> x=sqrt(t).*cos(t);
octave:3> y=sqrt(t).*sin(t);
octave:4> z=0.5*t;
octave:5> graph=plot3(x,y,z)
graph = -17.921
octave:6> set(graph(1), "linewidth", 2)
octave:7> 
```

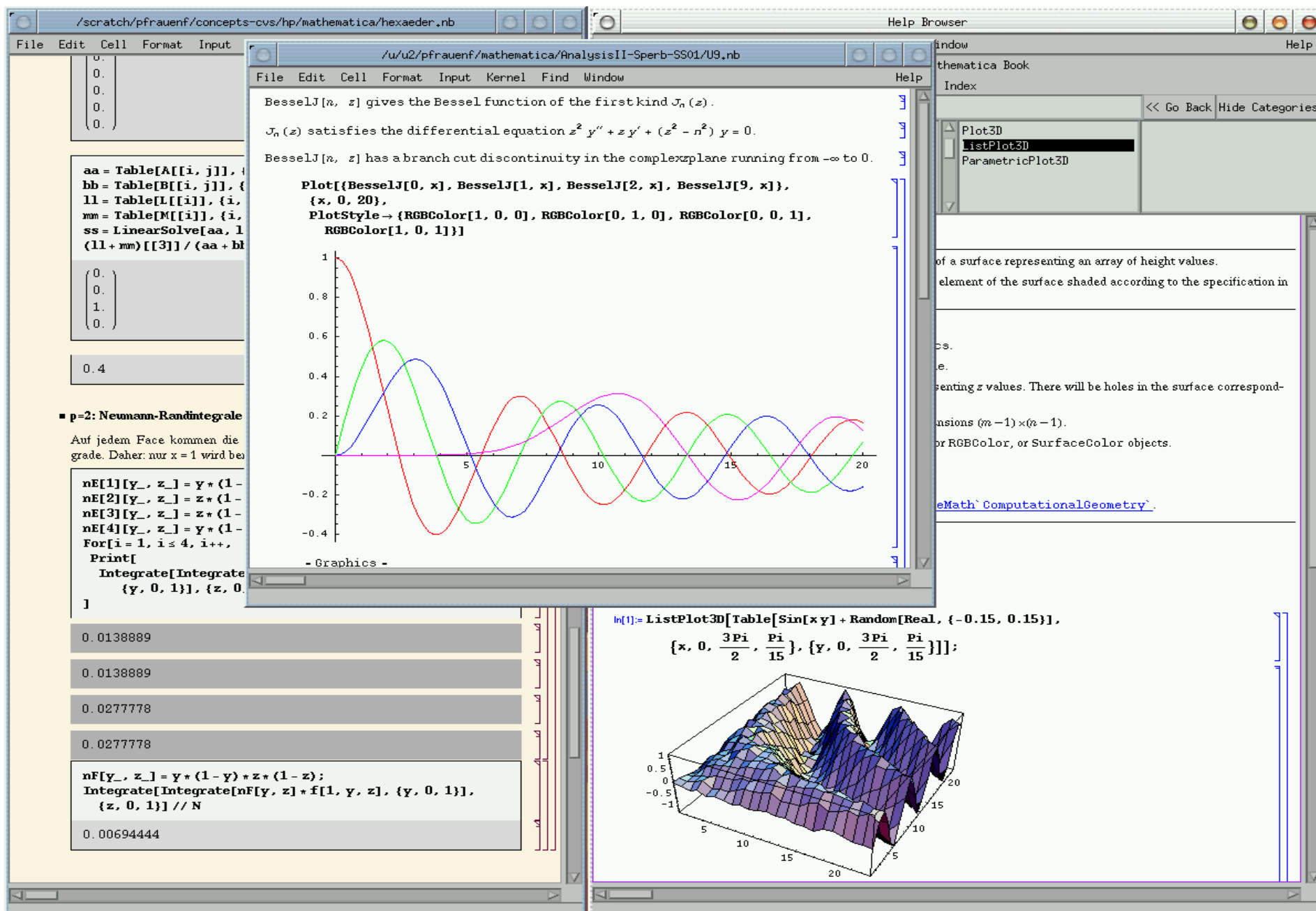

Interfejs Enthought Canopy



Obliczenia symboliczne – Mathematica i podobne

- Opracowana w latach 1980'tych przez Stephen'a Wolframa
- Jeden z pierwszych w historii pakietów umożliwiających obliczenia symboliczne
- Bardzo popularna wśród studentów amerykańskich, którzy muszą “zaliczyć” rachunek różniczkowy
- Obecnie także w wersji online: Wolfram Alpha
- Konkurencyjne pakiety: Maple, Mathcad, Symbolic math toolbox w matlabie (dawny muPAD)

Mathematica - interfejs



Wolfram Alpha – mathematica online



integrate sin x dx from x=0 to pi



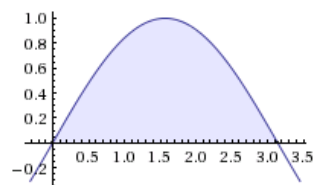
Examples Random

Definite integral:

Step-by-step solution

$$\int_0^{\pi} \sin(x) dx = 2$$

Visual representation of the integral:



Riemann sums:

More cases

left sum	$\frac{\pi \cot(\frac{\pi}{2n})}{n} = 2 - \frac{\pi^2}{6n^2} + O(\frac{1}{n^4})$
----------	--

(assuming subintervals of equal length)

Enable interactivity

cot(x) is the cotangent function

Indefinite integral:

Step-by-step solution

$$\int \sin(x) dx = -\cos(x) + \text{constant}$$

Download page

POWERED BY THE WOLFRAM LANGUAGE

Related Queries:

- y(x) - 3 (integrate y(z) sin(x+z) from z = 0...
- integrate using midpoint method sin(x) fr...
- integrate 1/sin(x) dx
- area between sinx and cosx from x = 0 t...
- area under y = sin(x) from x = 0 to pi

- Interfejs online umożliwiający korzystanie z wielu narzędzi do obliczeń symbolicznych
- Duża część funkcjonalności darmowa, ale wiele funkcji (np. rozpisywanie rozwiązań na kroki) - płatna

Obliczenia numeryczne a symboliczne

- W obliczeniach symbolicznych próbujemy obejść problem numerycznych zaokrągleń i przybliżeń poprzez opis równań algebraicznych explicite
- Tego typu pakiety pozwalają na dokładne odwzorowanie równań, jednak cierpią z powodu heurystycznych metod rozwiązania
- Proste operacje w takich pakietach są prostsze niż “na kartce” ale trudne często mogą nastroczać więcej problemów niż korzyści
- Na pewno są skuteczne do sprawdzania, czy nie pomyliliśmy się w obliczeniach

Typowe funkcje pakietów symbolicznych

- Przekształcanie, upraszczanie wzorów
- Rozwiązywanie równań i układów równań algebraicznych
- Znajdowanie granic wyrażeń i ciągów liczbowych
- Całkowanie i różniczkowanie symboliczne
- Wykresy
- Ładne formatowanie wzorów matematycznych (często przy użyciu LaTeX'a)

Maxima - Obliczenia symboliczne

Open Source

- Maxima (1992-), a wcześniej Macsyma (1968-1982)
- Wydana w 1998 na licencji GPL
- Napisana w języku lisp
- Wiele konkurencyjnych interfejsów (WXMaxima, Gmaxima itp)
- Skupiona na obliczeniach symbolicznych

Projekt SymPy

- Projekt narzędzi do obliczeń symbolicznych dla języka python
- Powstaje od ok. 2005 roku, obecnie osiągnął wersję 1.0
- Napisany w pythonie, kładzie nacisk na czytelność kodu i rozszerzalność, niekoniecznie na szybkość i pełność systemu
- Zawiera podstawowe funkcjonalności (zmienne symboliczne, granice, równania, różniczkowanie, całkowanie symboliczne)
- Dobrze integruje się z innymi pakietami w pythonie

Przykład użycia sympy

https://notch.mimuw.edu.pl:8888/notebooks/Symbolicznien.ipynb

Jupyter Symbolicznien Last Checkpoint: a few seconds ago (autosaved) Logout

File Edit View Insert Cell Kernel Help Python 2

Code CellToolbar

```
In [37]: import sympy, math
print math.sqrt(8), sympy.sqrt(8)
print math.sqrt(8)**2, sympy.sqrt(8)**2
sympy.Rational(1,3), 1./3

2.82842712475 2*sqrt(2)
8.0 8

Out[37]: (1/3, 0.3333333333333333)
```

```
In [26]: from sympy import symbols
x, y = symbols('x y')
expr = x + 2*y # wyrażenia symboliczne
print expr
print expr - x**2

x + 2*y
-x**2 + x + 2*y
```

```
In [30]: from sympy import expand, factor #wymnażanie
expanded_expr = expand(x*expr)
print x*expr, expanded_expr

x*(x + 2*y) x**2 + 2*x*y
```

```
In [31]: factor(expanded_expr) # grupowanie

Out[31]: x*(x + 2*y)
```

```
In [19]: sympy.diff(sin(x)*exp(x), x) # różniczkowanie

Out[19]: exp(x)*sin(x) + exp(x)*cos(x)
```

```
In [34]: sympy.integrate(exp(x)*sin(x) + exp(x)*cos(x), x) #całkowanie

Out[34]: exp(x)*sin(x)
```

```
In [35]: sympy.limit(sin(x)/x, x, 0) # granice

Out[35]: 1
```

```
In [22]: from sympy import latex
latex(Integral(cos(x)**2, (x, 0, pi))) # wypisywanie w LaTeX'u

Out[22]: '\\int_0^{\\pi} \\cos^2{\\left (x \\right )}\\, dx'
```

SAGE notebook

- Stosunkowo nowy projekt
- Połączenie wielu środowisk obliczeniowych
 - Python (Numpy, Scipy, SymPy, matplotlib, Networkx)
 - Maxima
 - R
 - GAP, FLINT, GD, JMOL, PALP, Singular
- Środowisko w przeglądarce, sesja na serwerze lub “w chmurze”

Interfejs SAGE

Use Sage to Solve Equations

last edited on April 11, 2011 05:45 PM by admin

[Save](#) [Save & quit](#) [Discard & quit](#)

File... Action... Data... **sage** ☐ Typeset

 [Print](#) [Worksheet](#) [Edit](#) [Text](#) [Undo](#) [Share](#) [Publish](#)

```
var('a b c d e f x y')
```

```
(a, b, c, d, e, f, x, y)
```

```
show(solve(a*x^2 + b*x + c == 0, x)[0])
```

$$x = -\frac{b + \sqrt{-4ac + b^2}}{2a}$$

```
show(solve(x^3 + a*x + b == 0, x)[0])
```

$$x = \frac{(-i\sqrt{3}+1)a}{6\left(\frac{1}{18}\sqrt{4a^3+27b^2}\sqrt{3}-\frac{1}{2}b\right)^{\frac{1}{3}}}-\frac{1}{2}(i\sqrt{3}+1)\left(\frac{1}{18}\sqrt{4a^3+27b^2}\sqrt{3}-\frac{1}{2}b\right)^{\frac{1}{3}}$$

```
solve([a*x + b*y == c, d*x + e*y == f], x, y)
```

```
[[x == -(b*f - c*e)/(a*e - b*d), y == (a*f - c*d)/(a*e - b*d)]]
```

[evaluate](#)

aaa - SageMathCloud - Mozilla Firefox

aaa - SageMathCloud

https://cloud.sagemath.com/projects/bc74: Search

Projects aaa Bartek Wilczynski About 133ms

Files New Log Find Settings 2016-05-30-002158.sagews

Modes Help # Data Control Program

x Plots Calculus Linear Graphs Number Theory Rings

1

2

3 `integrate(1 + x + x^2, x)`

4 $\frac{1}{3}x^3 + \frac{1}{2}x^2 + x$

5

6

7 `show(integrate(1 + x + x^2, x))`

8 `%var x, theta`

9
$$\frac{1}{3}x^3 + \frac{1}{2}x^2 + x$$

10

11 `solve(x**2-1,x)`

12

13 `[x == -1, x == 1]`

14

15

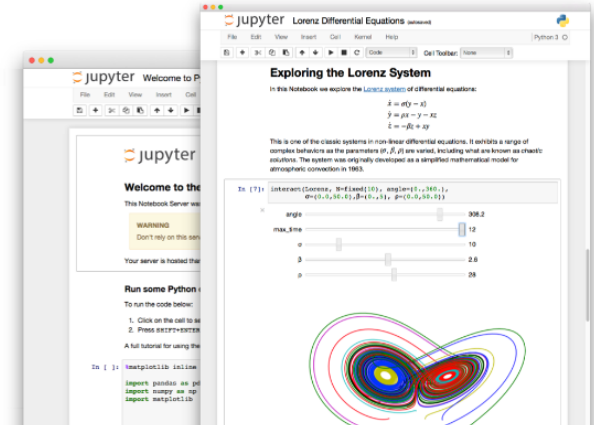
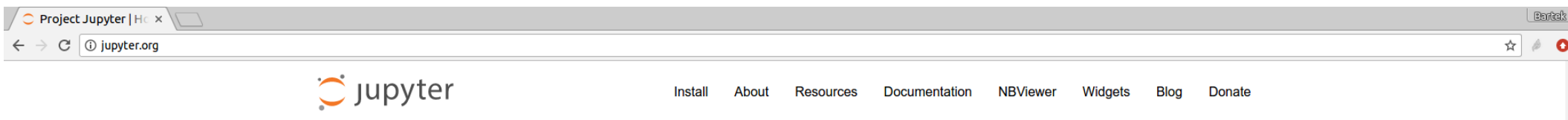
16 `plot(x**2-1)`

17

18

19

Jupyter notebook



The Jupyter Notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, machine learning and much more.



Language of choice

The Notebook has support for over 40 programming languages, including those popular in Data Science such as Python, R, Julia and Scala.



Share notebooks

Notebooks can be shared with others using email, Dropbox, GitHub and the [Jupyter Notebook Viewer](#).



Interactive widgets

Code can produce rich output such as images, videos, LaTeX, and JavaScript. Interactive widgets can be used to manipulate and visualize data in realtime.



Big data integration

Leverage big data tools, such as Apache Spark, from Python, R and Scala. Explore that same data with pandas, scikit-learn, ggplot2, dplyr, etc.

Excel?

- Najpopularniejszy pakiet do obliczeń
- Bardzo prosty interfejs
- Często stosowany również w bio-informatyce
- Ma spore ograniczenia (np. Maksymalna liczba linii w arkuszu), które utrudniają rozwój projektów prowadzonych w arkuszu
- Brak możliwości efektywnego testowania,
- Brak debuggerów
- Ma wiele funkcji, które warto znać, zwłaszcza, że często dane do obróbki dostajemy właśnie w Excel'u

COMMENT | OPEN ACCESS

Gene name errors are widespread in the scientific literature

Mark Ziemann, Yotam Eren and Assam El-Osta

Genome Biology 2016 17:177 | DOI: 10.1186/s13059-016-1044-7 | © The Author(s). 2016

Published: 23 August 2016

Abstract

The spreadsheet software Microsoft Excel, when used with default settings, is known to convert gene names to dates and floating-point numbers. A programmatic scan of leading genomics journals reveals that approximately one-fifth of papers with supplementary Excel gene lists contain erroneous gene name conversions.

Download PDF

Export citations

Table of Contents

Abstract
Declarations
References

METRICS

Article accesses: 62727

Citations: 5 more information

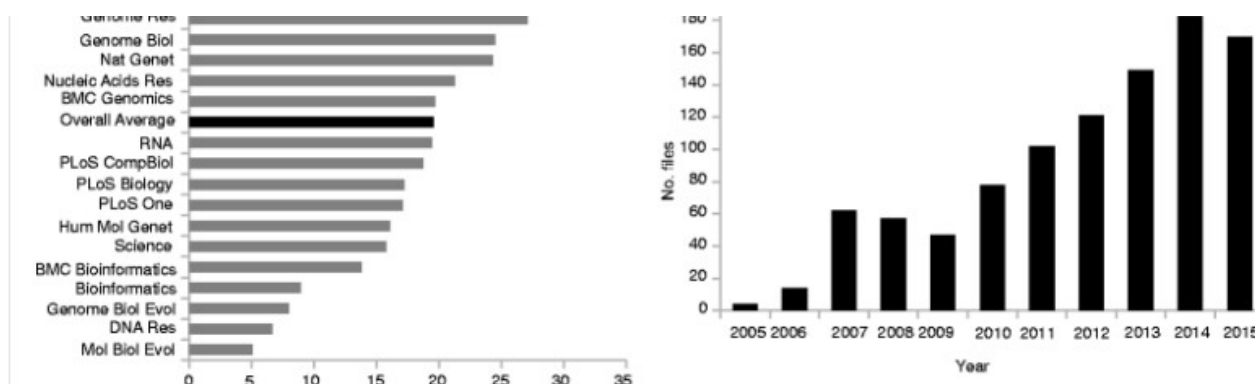


Fig. 1

Prevalence of gene name errors in supplementary Excel files. **a** Percentage of published papers with supplementary gene lists in Excel files affected by gene name errors. **b** Increase in gene name errors by year