

L^AT_EX e CSV

Massimiliano Dominici

Gruppo utilizzatori Italiani di T_EX

17 ottobre 2009

Piano della presentazione

- 1 CSV
- 2 datatool
 - Il pacchetto
 - Un esempio di mail merging
- 3 pgfplots
 - Il pacchetto
 - Un grafico da GNU Octave
- 4 Conclusioni

- 1 CSV
- 2 datatool
 - Il pacchetto
 - Un esempio di mail merging
- 3 pgfplots
 - Il pacchetto
 - Un grafico da GNU Octave
- 4 Conclusioni

Cosa significa CSV

- CSV = Comma Separated Values.
- I dati sono organizzati in tabelle, i cui *campi* sono separati da virgole.
- Formato **piatto**, ovvero tutti i dati sono inclusi in un'unica tabella.
- Non è uno standard, ma specifiche ufficiose sono indicate in RFC 4180.

Un esempio di CSV

```
Id,Autore,Titolo,Anno
001,Carlo Emilio Gadda,La Meccanica,1999
002,Ryunosuke Akutagawa,Rashomon e altri racconti,2008
003,Francesco De Sanctis,"Storia della
letteratura italiana",2006
004,Louis Stevenson,Weir di Hermiston,2000
005,Jan Tschichold,La forma del libro,2003
006,Matthew P. Shiel,La nube purpurea,2004
007,Giorgio Manganelli,Teatro,2008
008,Galileo Galilei,Opere,2005
009,Erodoto,Le storie,2006
```

Sintassi dei CSV

- Ogni *record* è separato da un carattere di fine riga (CRLF):

```
008,Galileo Galilei,Opere,2005 CRLF
009,Erodoto,Le storie,2006
```

- La prima riga può opzionalmente rappresentare un'intestazione contenente il nome dei campi:

```
Id,Autore,Titolo,Anno
001,Carlo Emilio Gadda,La Meccanica,1999
```

- L'intestazione e ciascun record possono contenere un numero qualsiasi di campi, separati da virgole:

```
Id,Autore,Titolo,Anno
001,Carlo Emilio Gadda,La Meccanica,1999
```

Sintassi dei CSV

- Ogni campo può essere opzionalmente racchiuso tra doppi apici (""). I doppi apici sono obbligatori se il campo contiene una virgola, un carattere di fine riga o i doppi apici:


```
"002",Ryunosuke Akutagawa,"Rashomon CRLF
e altri racconti",2008
```
- Se il campo contiene il carattere "", questo deve essere raddoppiato (e il campo racchiuso tra doppi apici).

Una grammatica formale del formato è contenuta in RFC 4180.

DSV: CSV “estesi”

Esistono altri formati con le stesse caratteristiche dei CSV, ma con diversi delimitatori di campo: “;”, “:”, “**TAB**”, “**SPACE**”.

Questi formati prendono il nome generico di:
Delimiter Separated Values (DSV).

Utilità dei CSV/DSV

- Varie forme di DSV sono usate come formato di scambio tra applicazioni diverse.
- Esportazione/importazione da fogli di calcolo (MS Excel, OpenOffice Calc, Gnumeric, ecc.).
- Esportazione di tabelle o viste da gestori di basi di dati (Oracle, MySQL, PostgreSQL, ecc.).
- Esportazione da programmi di calcolo numerico (Matlab, GNU Octave, Gnuplot, ecc.).

CSV/DSV e \LaTeX

- Gestione di piccole basi di dati o di rubriche per applicazioni da ufficio (report automatici, gestione della corrispondenza, ecc.).
- Pacchetto consigliato: **datatool**.
- Generazione di grafici con strumenti nativi di \LaTeX (PFG/TikZ) a partire da dati ottenuti da programmi esterni.
- Pacchetto consigliato: **pgfplots**.

- 1 CSV
- 2 datatool
 - Il pacchetto
 - Un esempio di mail merging
- 3 pgfplots
 - Il pacchetto
 - Un grafico da GNU Octave
- 4 Conclusioni

Presentazione del pacchetto

- Autore: Nicola Talbot.
- Funzioni: permette di creare, modificare, importare e visualizzare piccole basi di dati.
- Offre strumenti potenti per trasformare file CSV in tabelle di dati o per gestire rubriche di indirizzi.
- Non è adatto a gestire dati provenienti da programmi di calcolo numerico (non riconosce dati numerici in notazione scientifica).

Caricare un database esterno

- `\DTLloaddb [⟨opt⟩]{⟨db⟩}{⟨file⟩}`

Ad ogni campo viene assegnata una chiave identificativa, ricavata dall'intestazione. Le opzioni servono soprattutto nel caso nessuna intestazione sia presente.

noheader: nessuna intestazione presente

keys: assegna ai campi chiavi definite dall'utente

headers: assegna ai campi nomi definiti dall'utente

- `\DTLloadrawdb [⟨opt⟩]{⟨db⟩}{⟨file⟩}`

Traduce automaticamente $\$ \rightarrow \backslash \$$, $\% \rightarrow \backslash \%$, ecc.

Visualizzare un database esterno

```
\DTLloaddb{catalogo}{catalogo.csv}
\DTLdisplaydb{catalogo}
```

Id	Autore	Titolo	Anno
001	Carlo Emilio Gadda	La Meccanica	1999
002	Ryunosuke Akutagawa	Rashomon e altri racconti	2008
003	Francesco De Sanctis	Storia della letteratura italiana	2006
004	Louis Stevenson	Weir di Hermiston	2000
005	Jan Tschichold	La forma del libro	2003
006	Matthew P. Shiel	La nube purpurea	2004
007	Giorgio Manganelli	Teatro	2008
008	Galileo Galilei	Opere	2005
009	Erodoto	Le storie	2006

Visualizzare un database esterno (avanzato)

```

\DTLsort*{Anno}{catalogo}
\begin{tabular}{llr}
  \bfseries Autore & \bfseries Titolo & \bfseries Anno
  \DTLforeach*[\DTLisgt{\anno}{2000}]{catalogo}{%
    \aut=Autore,\tit=Titolo,\anno=Anno}{%
      \\
      \anno & \aut & \tit
    }
\end{tabular}

```

Anno	Autore	Titolo
2003	Jan Tschichold	La forma del libro
2004	Matthew P. Shiel	La nube purpurea
2005	Galileo Galilei	Opere
2006	Francesco De Sanctis	Storia della letteratura italiana
2006	Erodoto	Le storie
2008	Ryunosuke Akutagawa	Rashomon e altri racconti
2008	Giorgio Manganelli	Teatro

Un esempio di mail merging

Rubrica degli indirizzi dei soci `guit` indirizzi.csv:

Nome, Cognome, Via, Numero, Città, Prov, CAP
 Mario, Bianchi, Piazza Dante Alighieri, 7, Novi Ligure, AL, 15067
 Carlo, Rossi, Via Roma, 8, Noto, SR, 96017
 Franco, Verdi, Corso Garibaldi, 3, Senigallia, AN, 60017

guit — Gruppo Utilizzatori Italiani di T_EX

c/o Ufficio Statistica
 Scuola Superiore Sant'Anna
 Piazza Martiri della Libertà, 33
 56127 - Pisa

Pisa, 13 settembre 2009

a: MARIO BIANCHI
 Piazza Dante Alighieri, 7
 15067 - Novi Ligure (AL)

Oggetto: Adesione al Gruppo Utilizzatori Italiani di T_EX

Caro socio,

benvenuto nel `guit`. Ti siamo grati per la fiducia riposta nel nostro Gruppo: il tuo sostegno e la tua partecipazione rappresentano degli elementi determinanti per il supporto alla comunità L^AT_EX italiana e per il suo sviluppo.

Nell'invarti l'ultimo numero di `Aratexnica`, ti ricordiamo che puoi contribuire alla rivista inviando articoli, oppure aiutando nel lavoro di redazione, revisione e revisione dei testi. Maggiori dettagli sulle modalità sono disponibili all'indirizzo <http://www.guit.sasup.it/aratexnica.php>.

Ti porgo nuovamente il benvenuto e resto a disposizione per qualsiasi ulteriore informazione.

Cordialmente,

MASSIMILIANO DOMINICI

guit — Gruppo Utilizzatori Italiani di T_EX

c/o Ufficio Statistica
 Scuola Superiore Sant'Anna
 Piazza Martiri della Libertà, 33
 56127 - Pisa

Pisa, 13 settembre 2009

a: CARLO ROSSI
 Via Roma, 8
 96017 - Noto (SR)

Oggetto: Adesione al Gruppo Utilizzatori Italiani di T_EX

Caro socio,

benvenuto nel `guit`. Ti siamo grati per la fiducia riposta nel nostro Gruppo: il tuo sostegno e la tua partecipazione rappresentano degli elementi determinanti per il supporto alla comunità L^AT_EX italiana e per il suo sviluppo.

Nell'invarti l'ultimo numero di `Aratexnica`, ti ricordiamo che puoi contribuire alla rivista inviando articoli, oppure aiutando nel lavoro di redazione, revisione e revisione dei testi. Maggiori dettagli sulle modalità sono disponibili all'indirizzo <http://www.guit.sasup.it/aratexnica.php>.

Ti porgo nuovamente il benvenuto e resto a disposizione per qualsiasi ulteriore informazione.

Cordialmente,

MASSIMILIANO DOMINICI

Un esempio di mail merging

```

\DTLloaddb{indirizzi}{indirizzi.csv}

\DTLforeach*{indirizzi}{%
  \Nome=Nome, \Cognome=Cognome,%
  \Via=Via, \Numero=Numero,%
  \Citta=Città, \Prov=Prov, \CAP=CAP}{%
\begin{letter}{\Nome\ \Cognome\\
               \Via, \Numero\\
               \CAP\ -- \Citta\ (\Prov)}
\opening{Caro socio,}
<testo della lettera>
\closing{Cordialmente,}
\end{letter}
}

\signature{Massimiliano Dominici}
\city{Pisa}
\date{\today}
\subject{Adesione al \GuITtext}

```

- 1 CSV
- 2 datatool
 - Il pacchetto
 - Un esempio di mail merging
- 3 pgfplots
 - Il pacchetto
 - Un grafico da GNU Octave
- 4 Conclusioni

Presentazione del pacchetto

- Autore: Christian Feuersänger
- Funzioni: traccia grafici a partire da liste di coordinate specificate nel sorgente o caricate da un file esterno, oppure tramite l'indicazione di funzioni matematiche.
- Permette l'interazione con programmi di calcolo numerico esterni.
- Accetta valori numerici in notazione scientifica.
- Basato su PGF/TikZ, consente di personalizzare ogni aspetto visuale del grafico.

Comandi principali

- Ambiente `axis`: traccia gli assi delle coordinate e, opzionalmente, la griglia.
- Comando `\addplot`: traccia il grafico.
- Direttiva `plot file`: legge i dati da un file esterno, assumendo che siano fornite solo due colonne (in pratica ignora quelle in sovrannumero).
- Direttiva `plot table`: legge i dati da un file esterno, e consente di prelevarli da una colonna specificata dall'utente.

GNU Octave

- Programma per il calcolo numerico; equivalente Free Software di Matlab.
- Basato su una rappresentazione matriciale dei dati.
- Permette di esportare matrici di dati in forma tabulare in un file ASCII esterno.

Equazioni di Lotka-Volterra

- Modellano un sistema predatore-preda.
- Proposte indipendentemente da Alfred Lotka e Vito Volterra negli anni Venti.
- Sistema di due equazioni differenziali non lineari del primo ordine.

$$\begin{aligned}\frac{dx}{dt} &= (A - By)x \\ \frac{dy}{dt} &= (Cx - D)y\end{aligned}$$

Lo script di Octave

```
# PredPreyScript.m
t0 = 0;
tf = 20;
init_vals = [1; 0.5];
options=odeset('AbsTol',1.e-12,'RelTol',1.e-9,
  'InitialStep',2,'MaxStep',2);
[t,x] = ode45(@pred_prey_odes,[t0,tf],init_vals,options)
A = [t,x];
save -ascii PredPrey.dat A

#pred_prey_odes.m
function deriv_vals = pred_prey_odes(t,x)
deriv_vals = zeros(size(x));
deriv_vals(1) = 2*x(1) - 1.2*x(1).*x(2);
deriv_vals(2) = -1*x(2) + 0.9*x(1).*x(2);
```

Il file PredPrey.dat

```
0.00000000e+00 1.00000000e+00 5.00000000e-01
5.66993145e-03 1.00797044e+00 4.99726732e-01
4.51639430e-02 1.06532387e+00 4.98401027e-01
8.51085701e-02 1.12669546e+00 4.98125368e-01
1.25796018e-01 1.19282722e+00 4.99007906e-01
[...]
```

- Non è un CSV vero e proprio.
- Separatore di campo: uno o più spazi.
- Valori numerici in notazione scientifica.
- Manca l'intestazione.

Codice \LaTeX

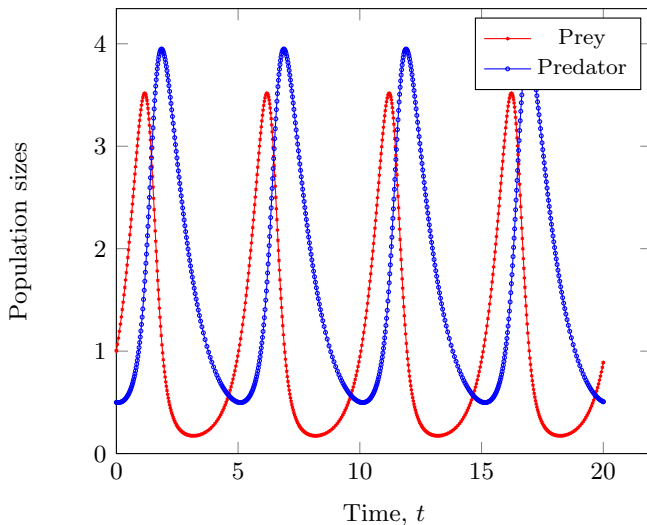
```

\begin{tikzpicture}
  \begin{axis}[ymin=0,xmin=0,
    xlabel={Time, $t$},
    ylabel={Population sizes},
    title={Our Predator Prey Example,
      Solution Over Time},
    legend entries={Prey,Predator}]
    \addplot[red,mark=asterisk,
      mark options={scale=.35}]
      plot table[x index=0,
        y index=1] {PredPrey.dat};
    \addplot[blue,mark=o,
      mark options={scale=.35}]
      plot table[x index=0,
        y index=2] {PredPrey.dat};
  \end{axis}
\end{tikzpicture}

```

Andamento nel tempo delle due popolazioni

Our Predator Prey Example, Solution Over Time



Codice L^AT_EX 2

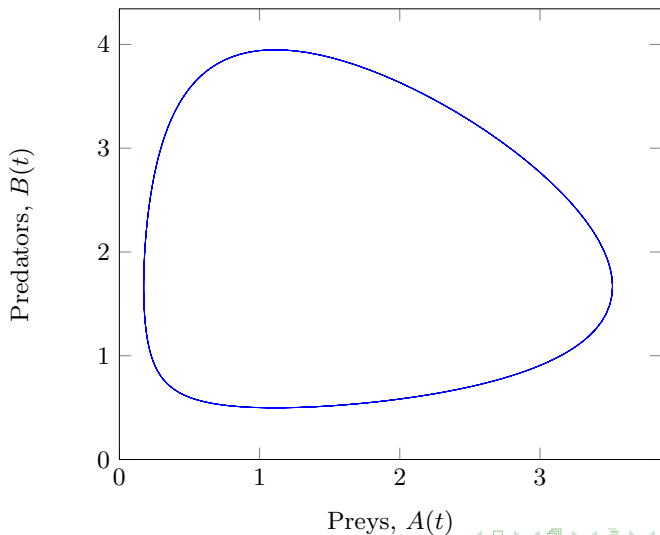
```

\begin{tikzpicture}
  \begin{axis}[ymin=0,xmin=0,
    xlabel={Preys,  $A(t)$ },
    ylabel={Predators,  $B(t)$ },
    title={Phase Plane for Predator Prey Example},
    line width=.1pt]
    \addplot[blue,mark=none]
      table[x index=1,y index=2] {PredPrey.dat};
  \end{axis}
\end{tikzpicture}

```

Piano delle fasi

Phase Plane for Predator Prey Example



- 1 CSV
- 2 datatool
 - Il pacchetto
 - Un esempio di mail merging
- 3 pgfplots
 - Il pacchetto
 - Un grafico da GNU Octave
- 4 Conclusioni

Altri usi di datatool e pgfplots

- **datatool** può essere usato anche per generare semplici grafici, istogrammi o diagrammi circolari.
- **pgfplots** può essere usato anche per visualizzare tabelle di dati, principalmente numerici.

Strategie consigliabili

- Organizzare bene i dati al momento dell'esportazione:
 - “appiattare” la struttura dei dati;
 - applicare filtri e ordinamenti ogni volta che è possibile.
- Usare **datatool** per “applicazioni da ufficio”:
 - semplici report statistici;
 - gestione di rubriche di indirizzi.
- Usare **pgfplots** per il tracciamento di grafici a partire da valori numerici ottenuti con programmi di calcolo esterni.

Grazie per l'attenzione.