



Kepler Exercise

Deana Pennington
University of New Mexico
December 10, 2004





Graphing data

Vegetation biomass test data

1. With EML
2. Without EML





Sevilleta LTER Biomass Plots

Biomass = $f(\text{species, cover\%, height, season})$

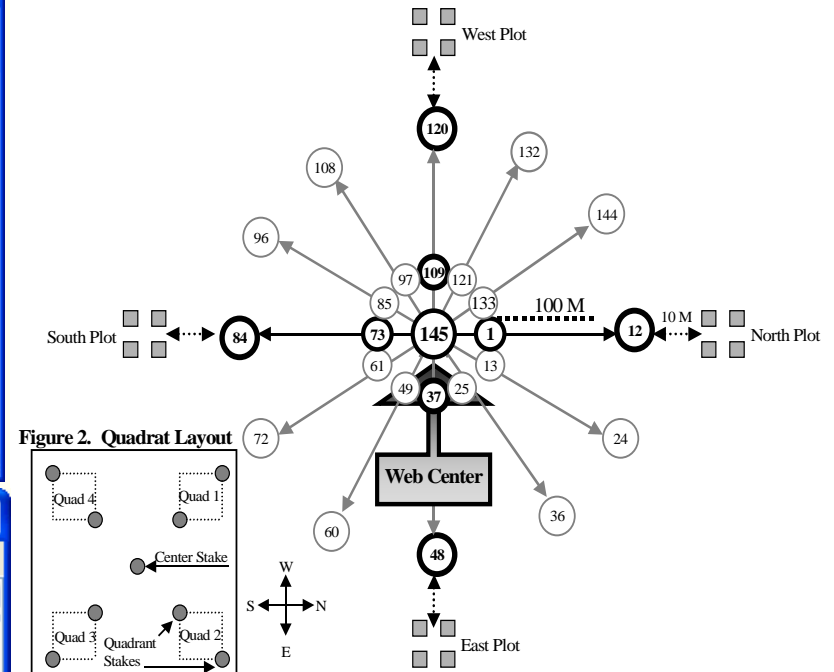
Microsoft Excel - 1999_npp.csv

	A	B	C	D	E	F	G	H	I	J	K	L
1	DATE	SITE	WEB	PLOT	QD	SPECIES	OBS	COVER	HEIGHT	COUNT	PHEN	COMMENT
2	2/3/1999	FPC	1	E	1	ERPU8	1	0.5	4	13	V	NA
3	2/3/1999	FPC	1	E	1	ERPU8	2	0.1	2	16	V	NA
4	2/3/1999	FPC	1	E	1	GUSA2	1	0.01	4	2	V	NA
5	2/3/1999	FPC	1	E	1	GUSA2	2	0.1	5	1	V	NA
6	2/3/1999	FPC	1	E	1	GUSA2	3	0.5	12	1	V	NA
7	2/3/1999	FPC	1	E	1	LEFE	1	0.25	5	1	V	NA
8	2/3/1999	FPC	1	E	2	LATR2	1	7	36	2	V	NA
9	2/3/1999	FPC	1	E	2	LATR2	2	2	32	3	V	NA
10	2/3/1999	FPC	1	E	2	LATR2	3	8	61	1	V	NA
11	2/3/1999	FPC	1	E	2	LATR2	4	3	45	1	V	NA
12	2/3/1999	FPC	1	E	2	LATR2	5	2	24	1	V	NA
13	2/3/1999	FPC	1	E	2	ERPU8	1	0.25	3	3	V	NA
14	2/3/1999	FPC	1	E	2	ERPU8	2	0.05	2	11	V	NA
15	2/3/1999	FPC	1	E	2	LEFE	1	0.1	4	3	V	NA

1999_npp.csv - Notepad

```

File Edit Format View Help
DATE, SITE, WEB, PLOT, QD, SPECIES, OBS, COVER, HEIGHT, COUNT, PHEN, COMMENTS
2/3/1999, FPC, 1, E, 1, ERPU8, 1, 0.5, 4, 13, V, NA
2/3/1999, FPC, 1, E, 1, ERPU8, 2, 0.1, 2, 16, V, NA
2/3/1999, FPC, 1, E, 1, GUSA2, 1, 0.01, 4, 2, V, NA
2/3/1999, FPC, 1, E, 1, GUSA2, 2, 0.1, 5, 1, V, NA
2/3/1999, FPC, 1, E, 1, GUSA2, 3, 0.5, 12, 1, V, NA
2/3/1999, FPC, 1, E, 1, LEFE, 1, 0.25, 5, 1, V, NA
2/3/1999, FPC, 1, E, 2, LATR2, 1, 7, 36, 2, V, NA
2/3/1999, FPC, 1, E, 2, LATR2, 2, 2, 32, 3, V, NA
2/3/1999, FPC, 1, E, 2, LATR2, 3, 8, 61, 1, V, NA
2/3/1999, FPC, 1, E, 2, LATR2, 4, 3, 45, 1, V, NA
2/3/1999, FPC, 1, E, 2, LATR2, 5, 2, 24, 1, V, NA
2/3/1999, FPC, 1, E, 2, ERPU8, 1, 0.25, 3, 3, V, NA
  
```



This Exercise

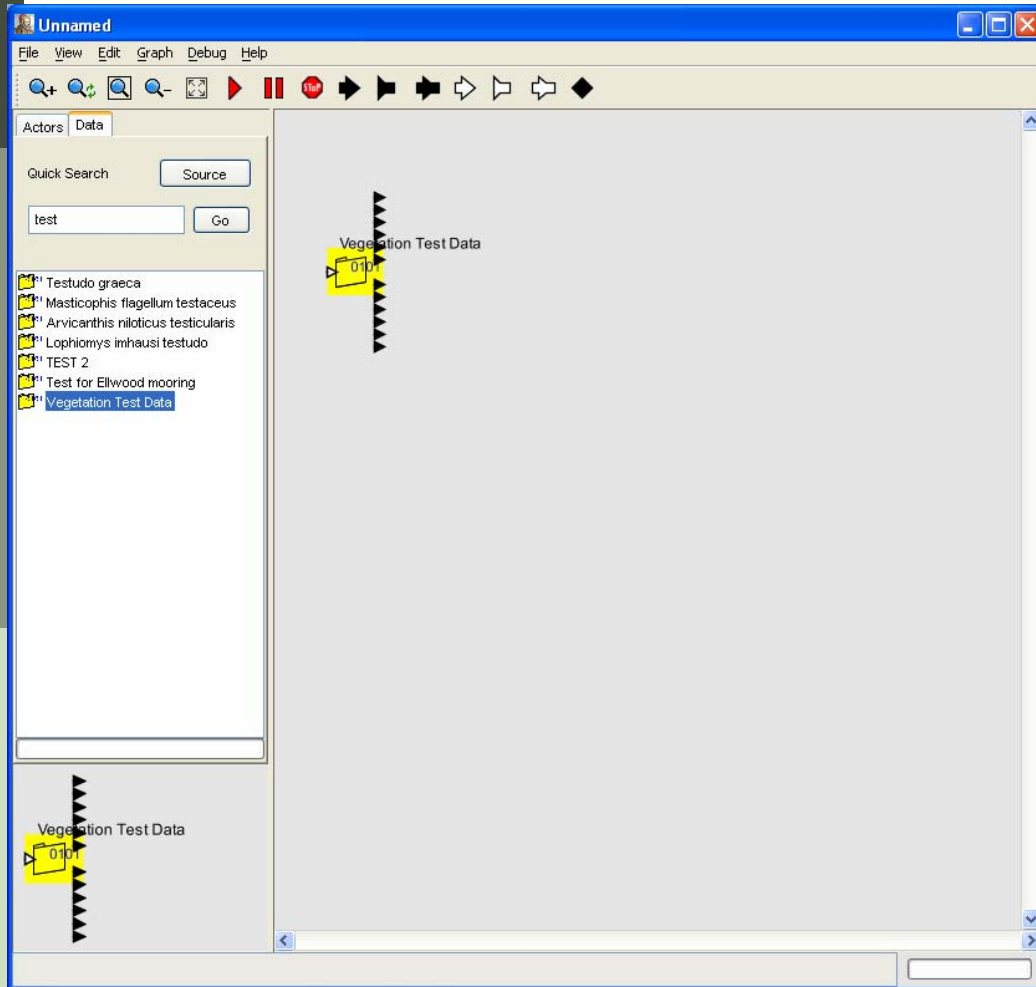
- Using a subset of the data for 1 species in 1 quadrat, extract vegetation cover% and height, and create bar graphs of those values through time

```

veg_data_test.csv - Notepad
File Edit Format View Help
DATE,SITE,WEB,PLOT,QD,SPECIES,OBS,COVER,HEIGHT,COUNT,PHEN,COMMENTS
2/3/1999,FPC,1,E,1,ERPU8,1,0.5,4,13,V,NA
2/3/1999,FPC,1,E,1,ERPU8,2,0.1,2,16,V,NA
6/2/1999,FPC,1,E,1,ERPU8,1,0.5,6,2,NA,NA
6/2/1999,FPC,1,E,1,ERPU8,2,0.25,4,12,NA,NA
6/2/1999,FPC,1,E,1,ERPU8,3,0.1,3,10,NA,NA
6/2/1999,FPC,1,E,1,ERPU8,4,0.05,2,13,NA,NA
10/7/1999,FPC,1,E,1,ERPU8,1,0.25,7,5,F,NA
10/7/1999,FPC,1,E,1,ERPU8,2,0.1,7,2,F,NA
10/7/1999,FPC,1,E,1,ERPU8,3,0.01,2,31,F,NA
    
```



With metadata (easier)



- ❑ Open a new graph editor in Kepler
File->New->Graph Editor
- ❑ Switch to the Data tab
- ❑ Search for "test"
- ❑ Add Vegetation Test Data to canvas
Drag and drop
- ❑ Display ports
Move the icon slightly to refresh





Explore the dataset

Edit parameters for Vegetation Test Data

EML File: Browse

schemaDef:

```
<schema>
  <table name="veg_data_test.csv">
    <field name="DATE" dataType="STRING"/>
    <field name="SITE" dataType="STRING"/>
    <field name="WEB" dataType="STRING"/>
    <field name="PLOT" dataType="STRING"/>
    <field name="QD" dataType="STRING"/>
    <field name="SPECIES" dataType="STRING"/>
    <field name="OBS" dataType="STRING"/>
    <field name="COVER" dataType="STRING"/>
  </table>
</schema>
```

sqlDef:

outputType:

Target File Extension in Compressed File:

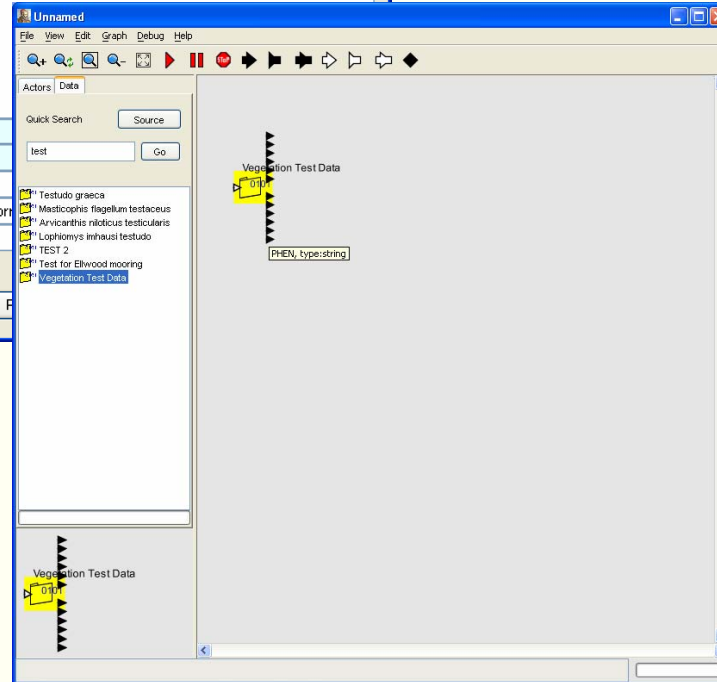
recordId:

endpoint:

veg_data_test-csv:

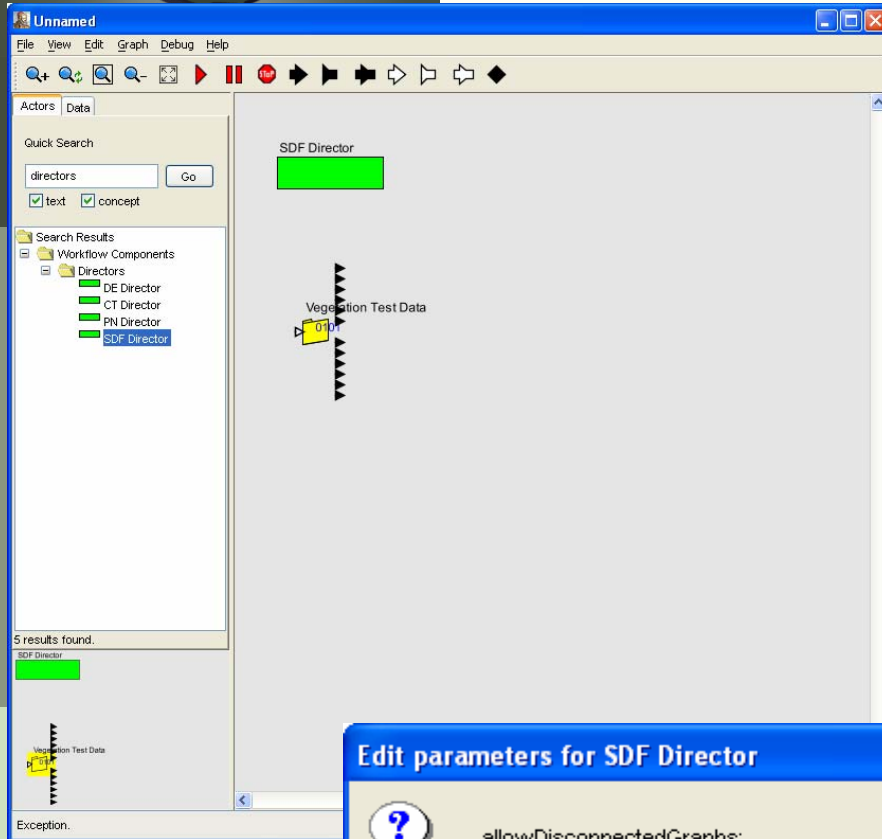
Commit Add Remove Restore Defaults

- Show the schema
Double-click on icon or
right-click-
>configure
- Tool tips on ports

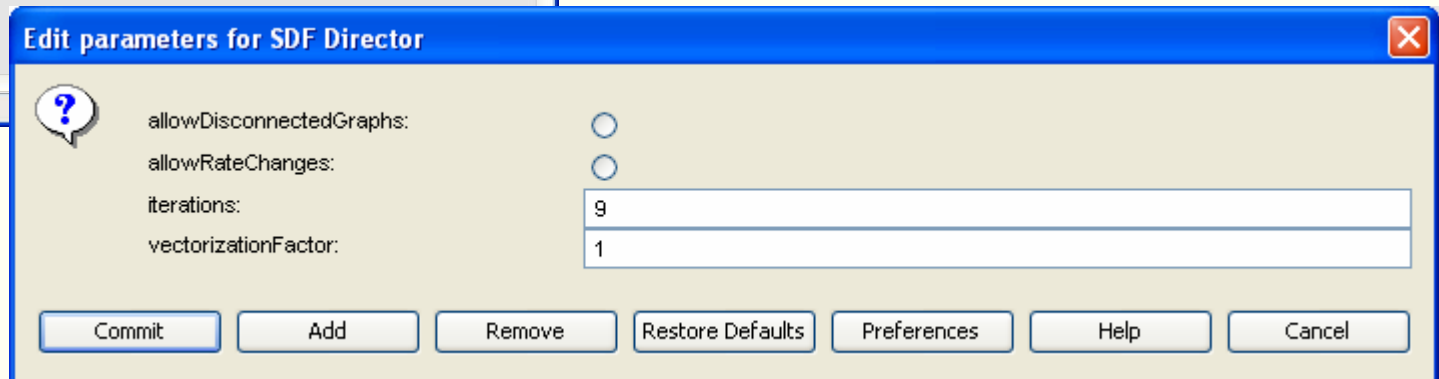




Director

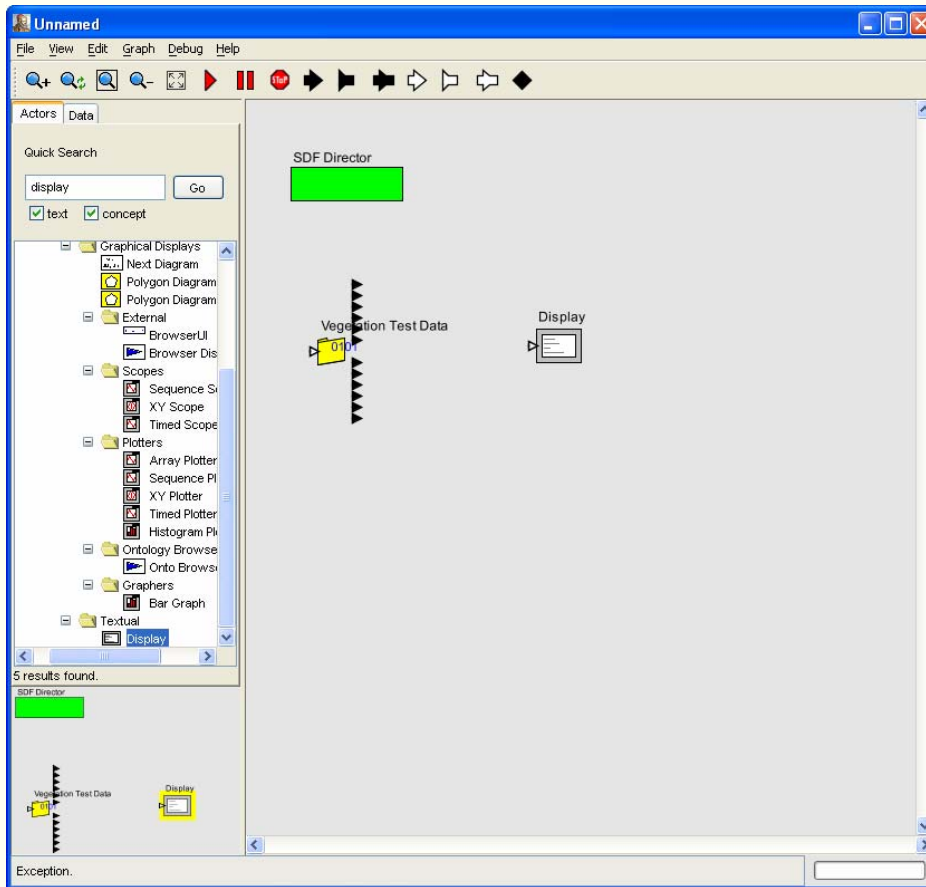


- ❑ Add a director
 - Click on the Actors tab
 - Double-click on Workflow Components
 - Double-click on Directors
 - Drag and drop the SDF Director onto the canvas
- ❑ Configure director
 - Double-click or right-click->configure
 - Change iterations to 9 (the number of rows to be read in the file)
 - Click Commit to save





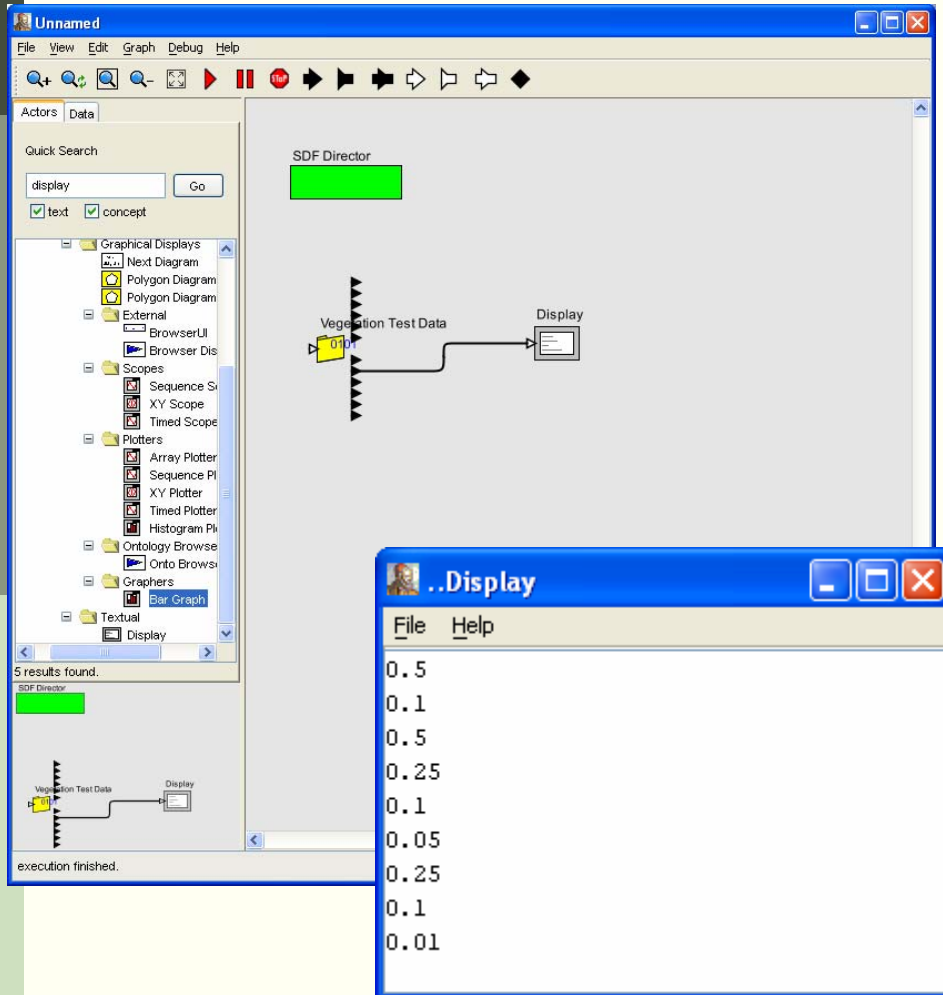
Display the file



- ❑ Search for “display”
- ❑ Find the Display actor
- ❑ Drag and drop the Display actor onto the canvas



Hooking up ports

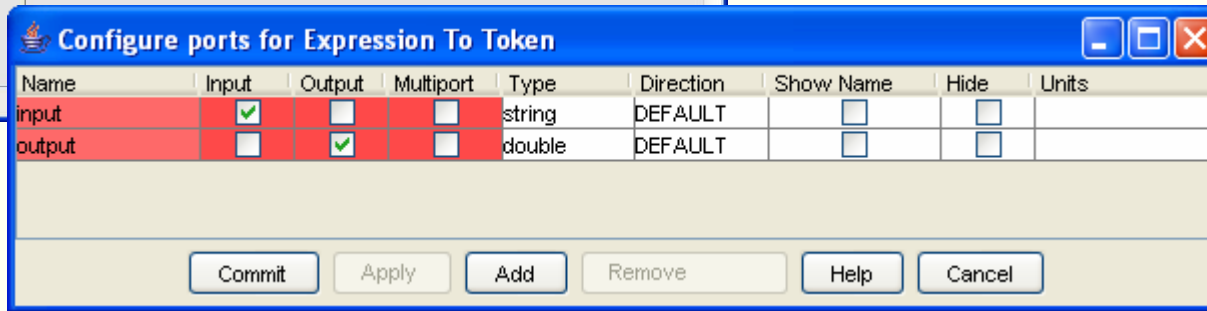
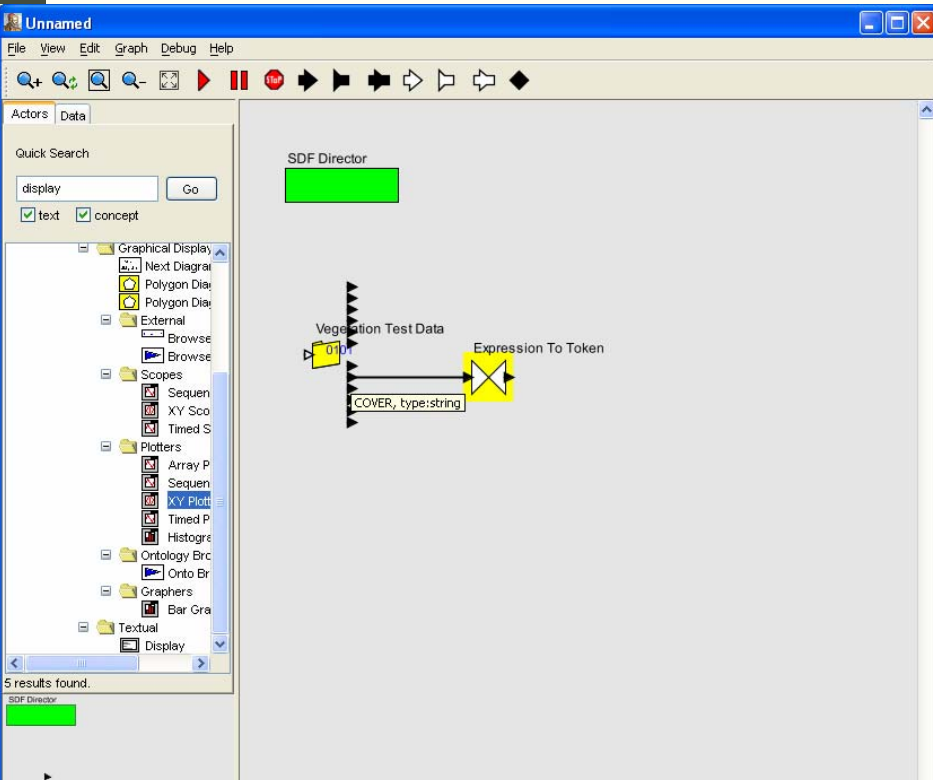


- ❑ Find the "cover" port
 - Use tool tips or view schema
 - Note string types
- ❑ Click-hold the cover port and drag to the input port on the Display actor
- ❑ Run the workflow
 - Click the red triangle or View->Run Window->Go
- ❑ Delete the Display actor



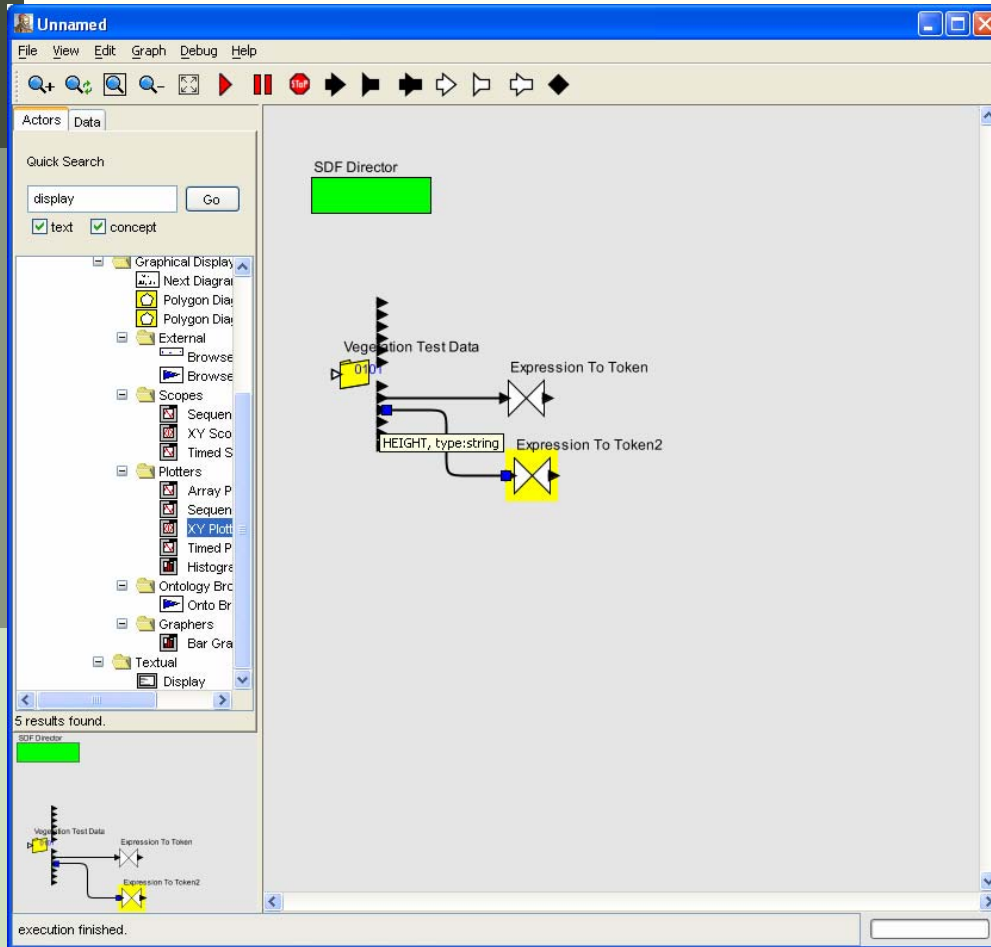
String conversion

- ❑ Search for Expression to Token actor
- ❑ Add Expression to Token actor to canvas
- ❑ Configure ports
 - Right-click->configure ports
 - Change input port Type to string
 - Change output port Type to double
 - Commit
- ❑ Hook up Cover data output port with Expression to Token input port





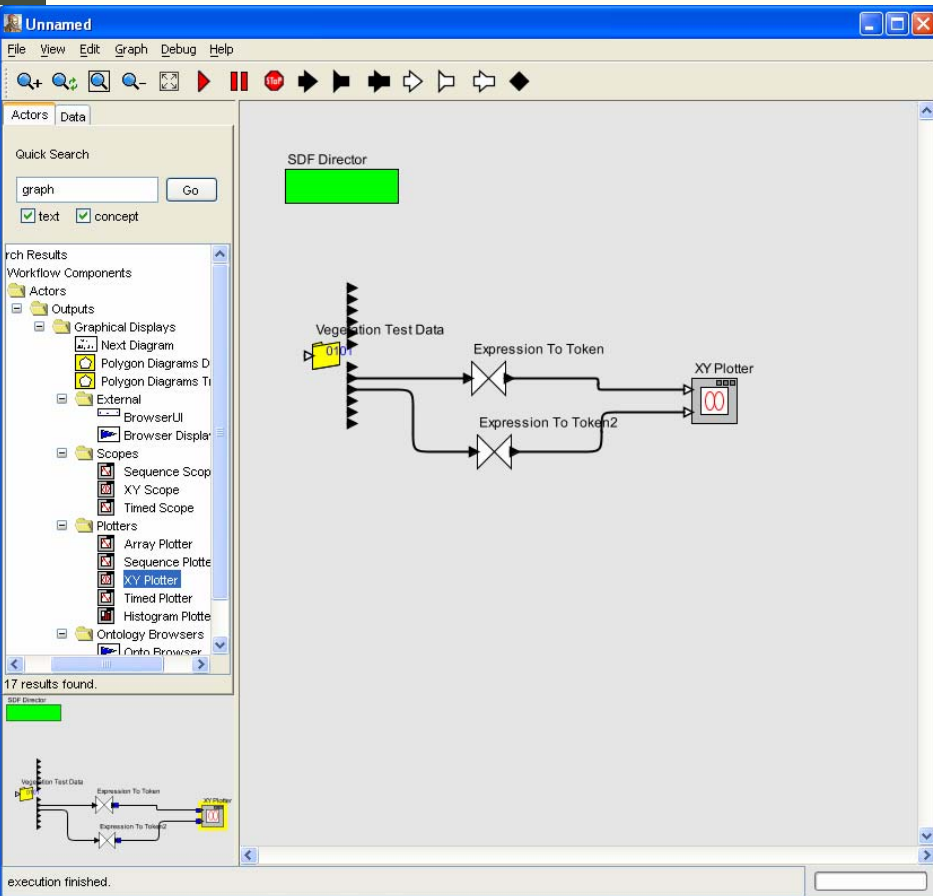
String conversion



- Repeat for Height data output port
- Copy and paste Expression to Token actor (Edit-> Copy-> Paste or cntl-c cntl-v), Move copy from over original
- Hook up height data output port with Expression to Token input port



Add graph

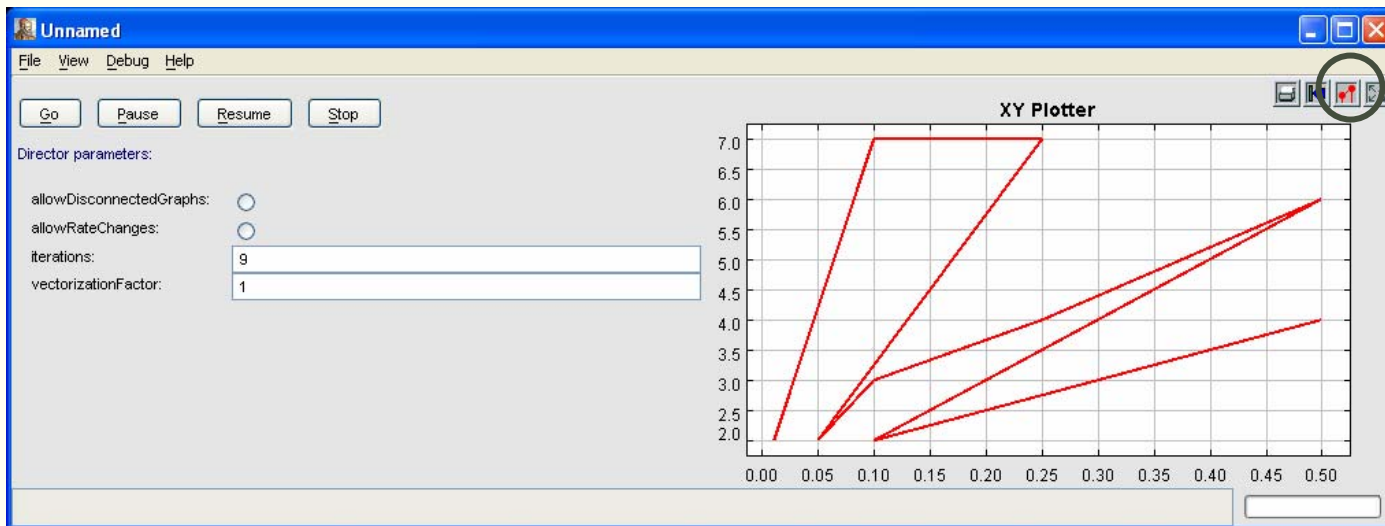


- ❑ Search for "xy plotter"
- ❑ Add XY Plotter actor to canvas
- ❑ Hook up Expression to Token output ports with XY Plotter input ports



Run the workflow

- ❑ Click the Run button (red triangle)
- ❑ Note lousy plot format..close the display
- ❑ Change the plot format
View->Run Window
Click on small plot format icon





Reformat the plot

- ❑ Change Title to something meaningful (I used Vegetation Volume)
- ❑ Add axis labels
- ❑ Select dots
- ❑ Unselect Connect
- ❑ Apply
- ❑ Close Run Window
- ❑ Run from red triangle
- ❑ Save your work

File -> Save ->

c:\kepler\workflows\eco\
veg_data_xxx.xml,
where xxx is your initials

Set plot format

? Title: Vegetation volume

X Label: Cover percent

Y Label: Height

X Range: 0.01, 0.5

Y Range: 2.0, 7.0

Marks: ☐ none ☐ points ☒ dots ☐ various ☐ pixels

Grid: ☒

Stems: ☐

Connect: ☐

Use Color: ☒

X Ticks:

Y Ticks:

Apply Cancel





Without EML

- ❑ The harder way, without EML-described metadata



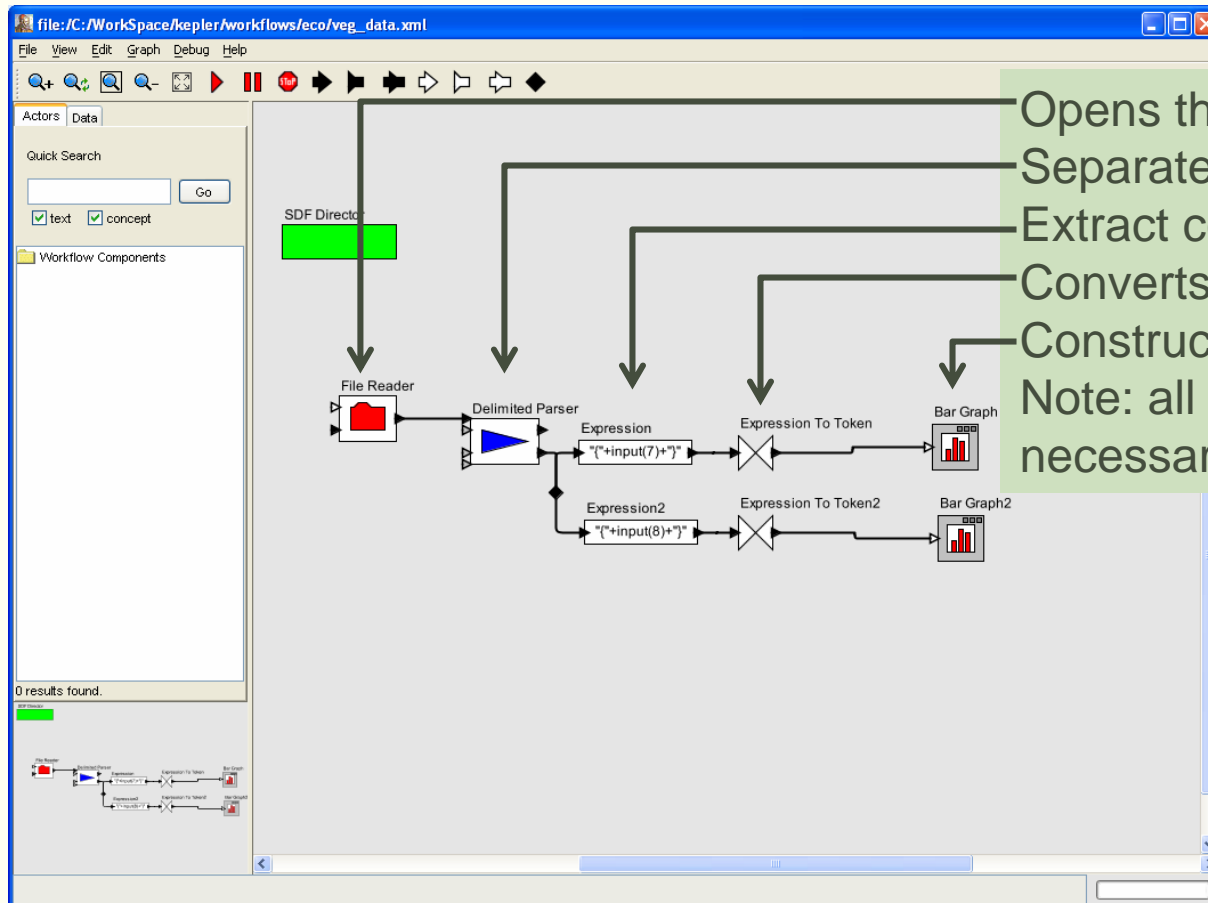


Cheat-look ahead

Copy veg_data_test.csv from shared directory to
c:\kepler\workflows\eco\

Copy veg_data.xml from share directory to c:\kepler\workflows\eco\

Open veg_data.xml from Kepler->File->Open File-> *[locate file]*



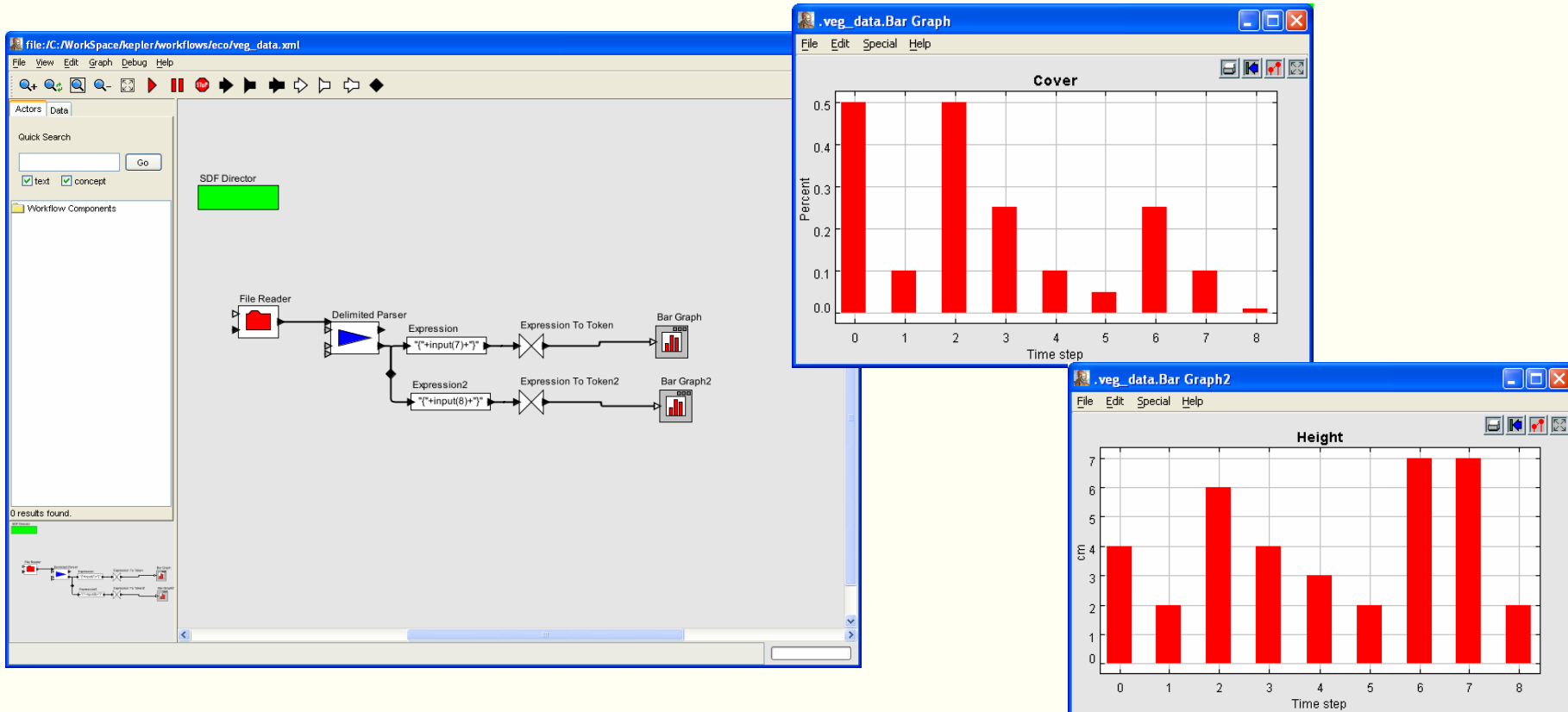
Opens the file
Separates file into rows/columns
Extract cover and height columns
Converts to decimal number
Constructs bar graphs
Note: all this would not be necessary if EML used





Play with the workflow

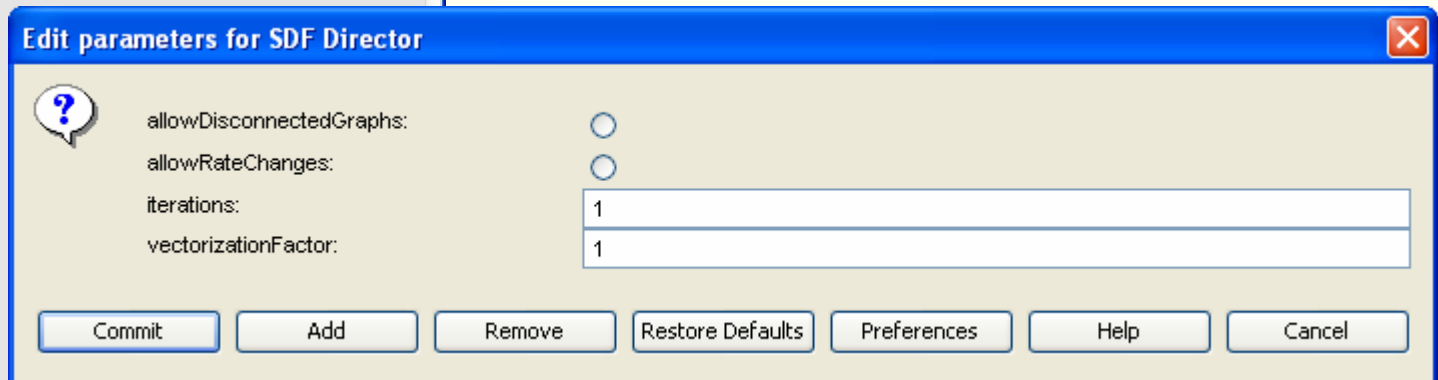
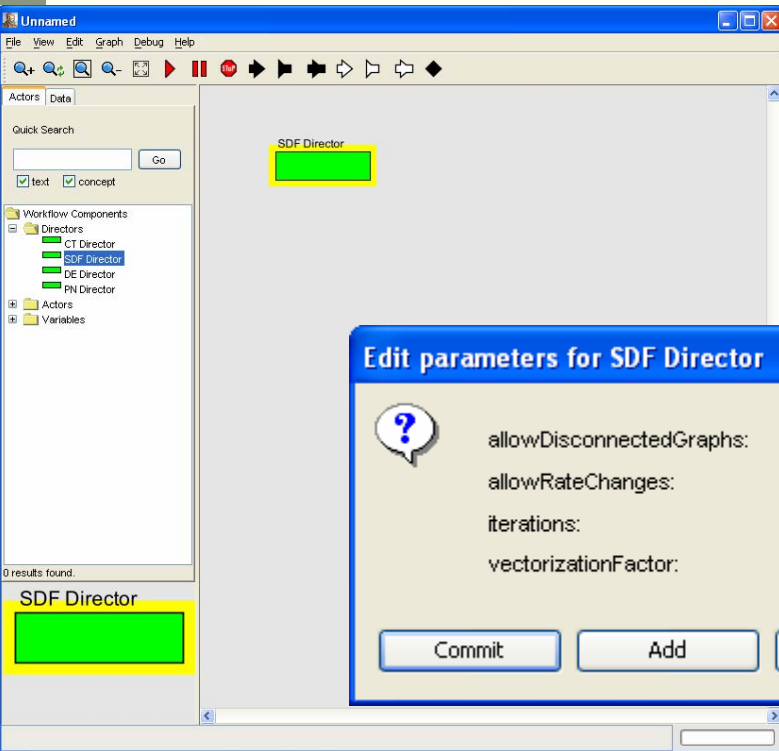
- ❑ Run workflow (click the red triangle)
- ❑ Move the top graph so both can be viewed
- ❑ Animate the execution Debug->Animate execution – set for 1000 ms
- ❑ Run workflow...move graph windows so you can see the animation
- ❑ Close the bar graphs
- ❑ Run from run window - View->Run Window-> Go





Start a new workflow

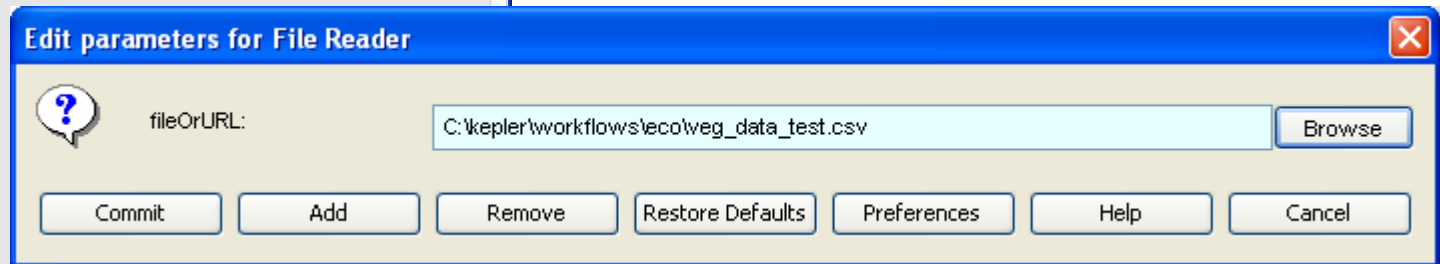
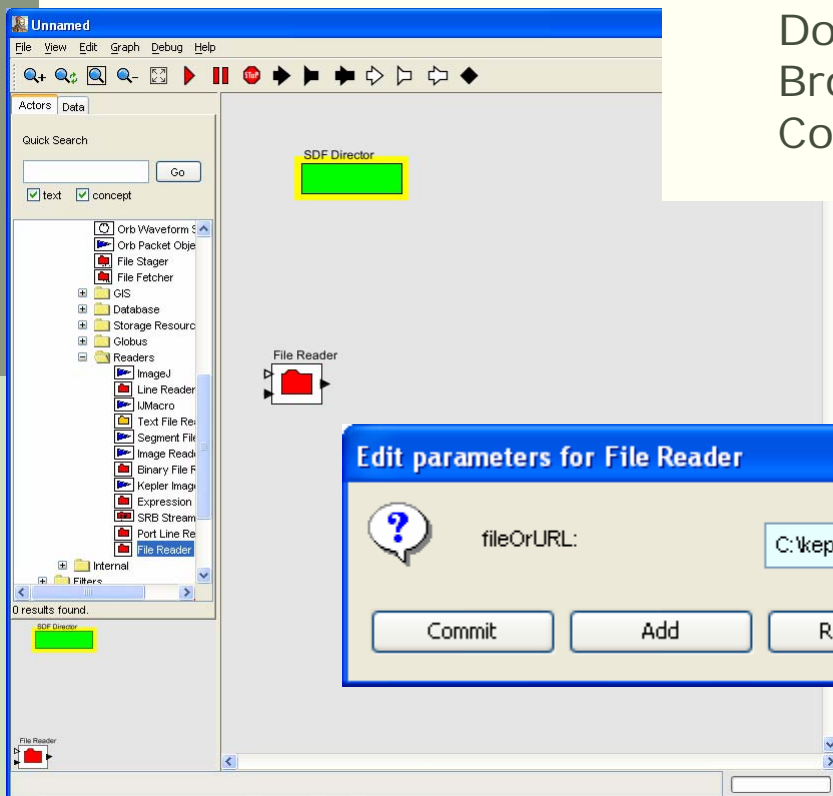
- ❑ Open the workflow editor
 - File -> New -> Graph Editor
- ❑ Select a director
 - Double click Workflow Components
 - Double click Directors
 - Drag and drop SDF Director
- ❑ Configure director
 - Right click on director
 - Change number of iterations from 0 to 1
 - Commit





Read the file

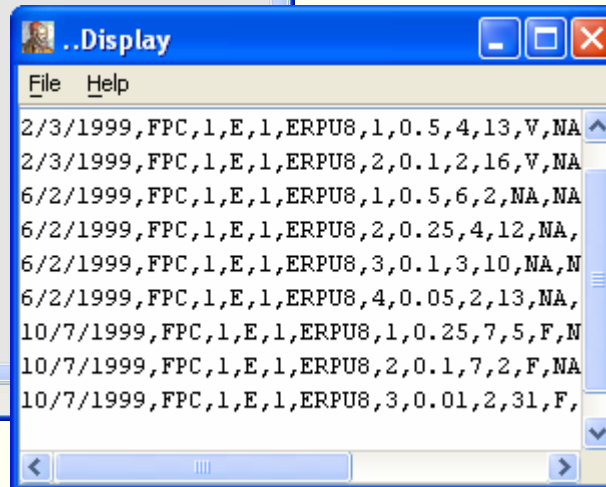
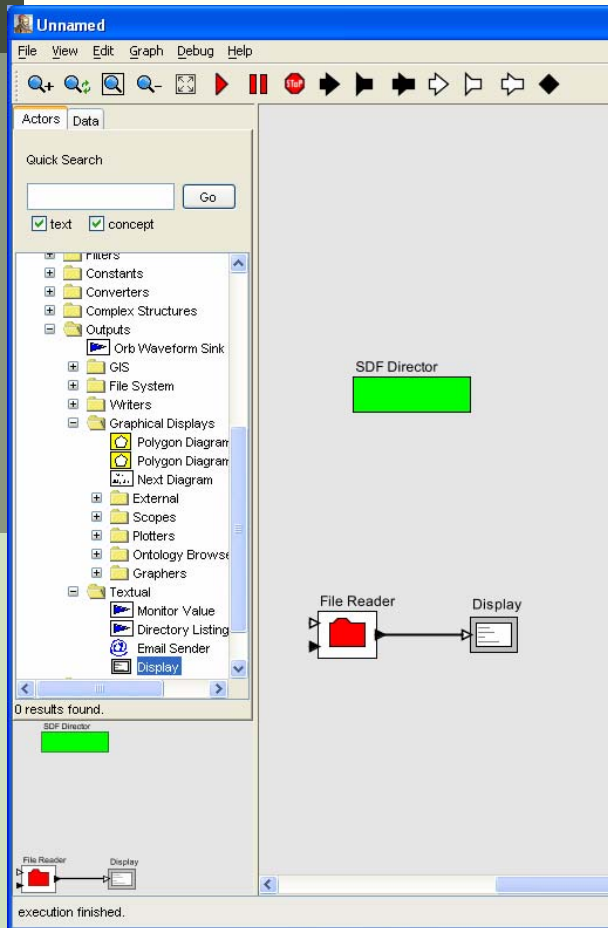
- ❑ Add the File Reader actor to the canvas
Select Actors->Inputs->External->Readers
Drag and drop the File Reader actor
- ❑ Configure the actor
Double-click or right-click->configure
Browse for the veg_data_test.csv file
Commit





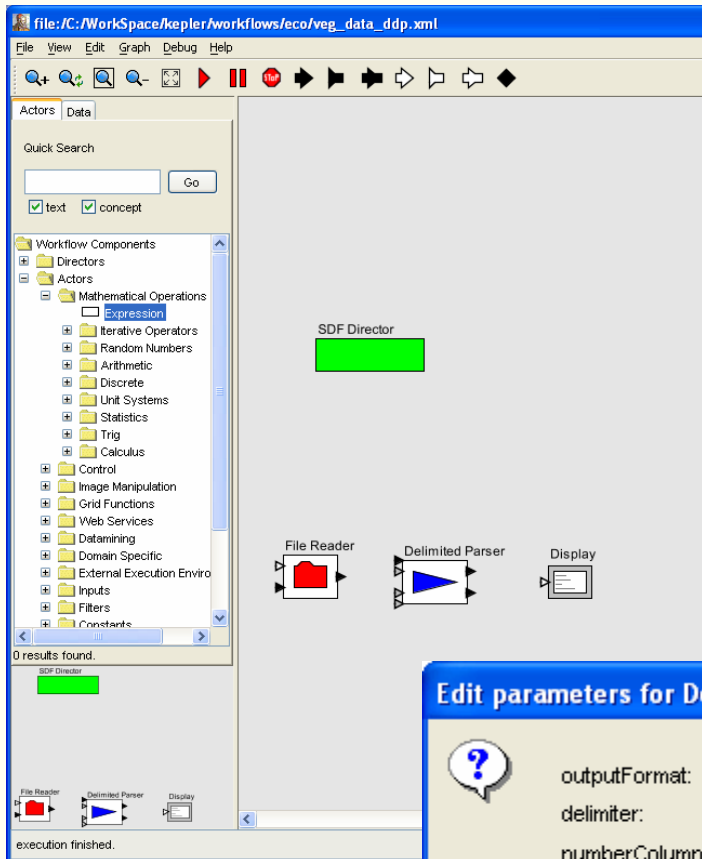
Display to check

- ❑ Add a Display actor, to check that the actor is working
 - Search for “display”
 - Drag and drop the Display actor
- ❑ Hook up ports between the 2 actors
 - Click-hold the output (right) port on the File Reader actor and drag to the input (left) port on the display actor, then release
- ❑ Run the workflow

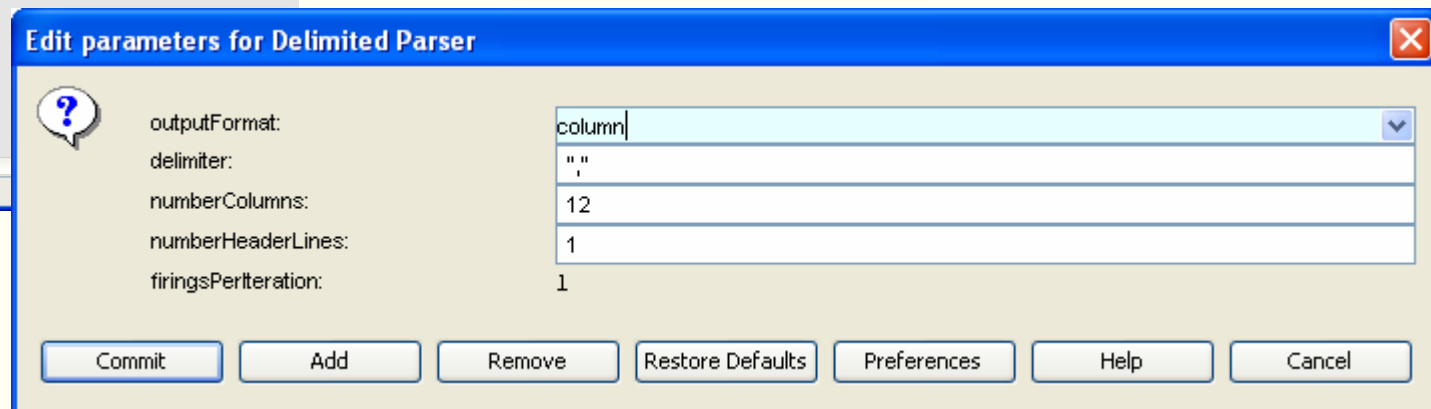




Parse the file

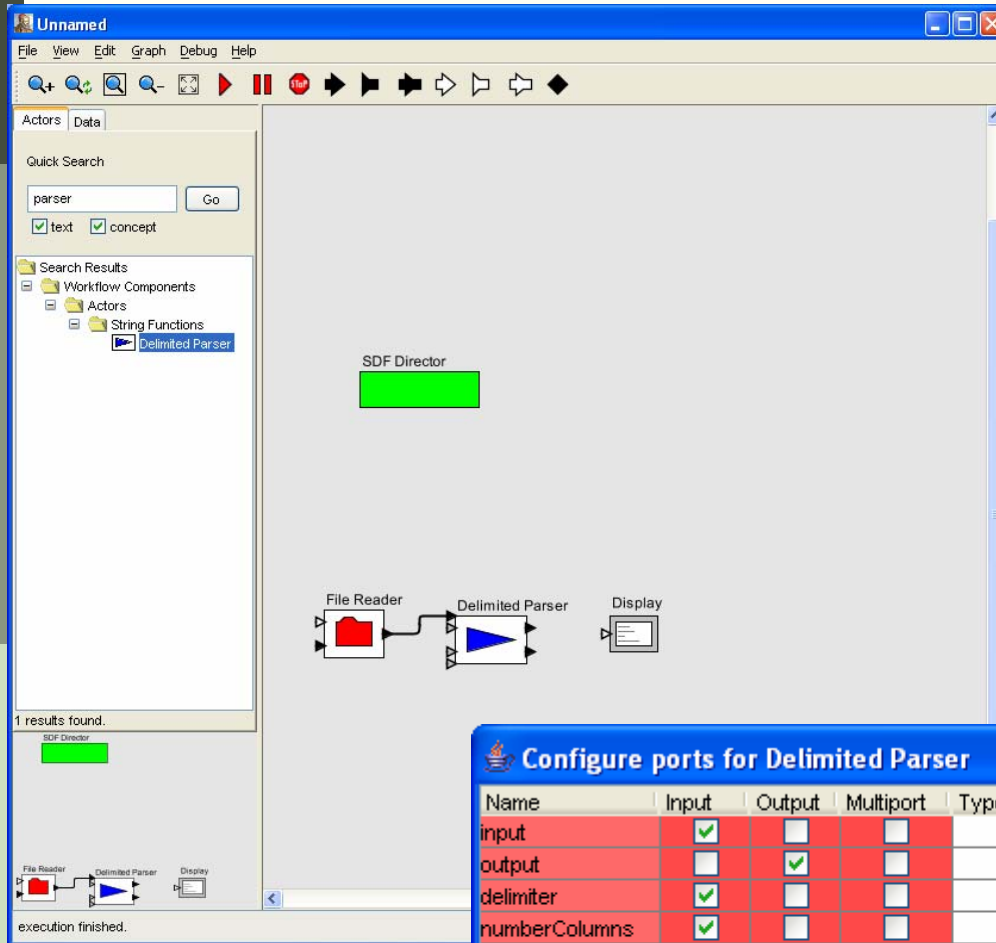


- Move (but keep) Display actor
Delete port connection (click->delete)
Move actor to right
- Quick Search for Delimited Parser actor
Type "parser" in text entry box
Click Go
- Drag and drop Delimited Parser onto canvas adjacent to File Reader
- Configure Delimited Parser
Double-click or right-click->Configure
outputFormat = column
Delimiter = ","
numberColumns = 12 (refer to Display actor output)
numberHeaderLines = 1 (refer to Display actor output)
CCommit

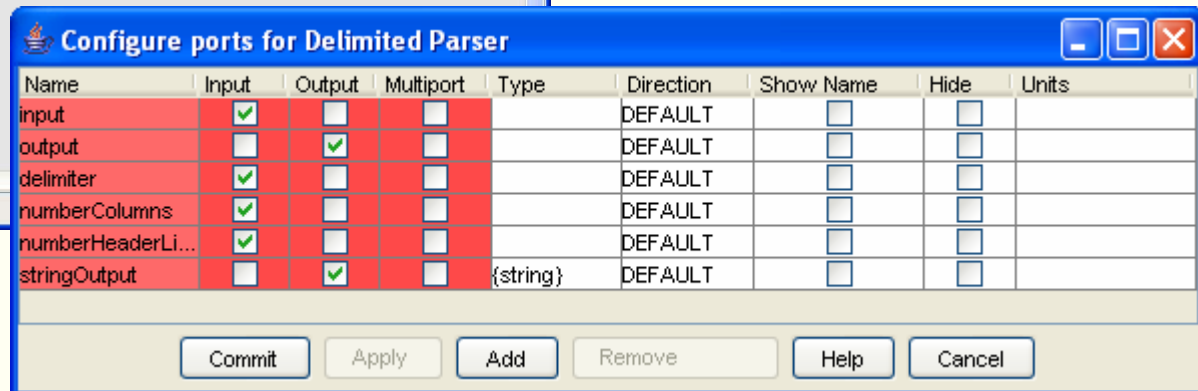




Parse the file (cont'd)



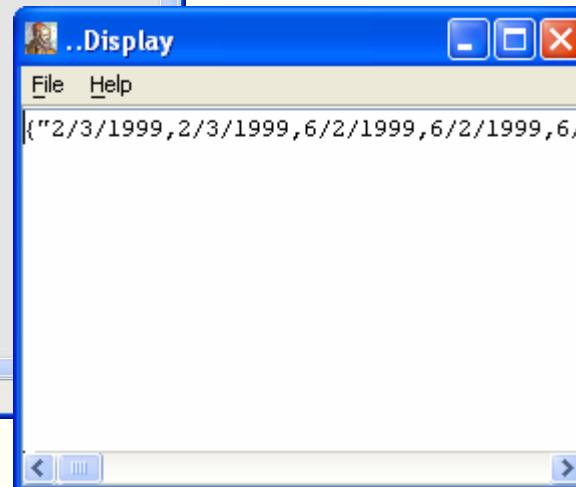
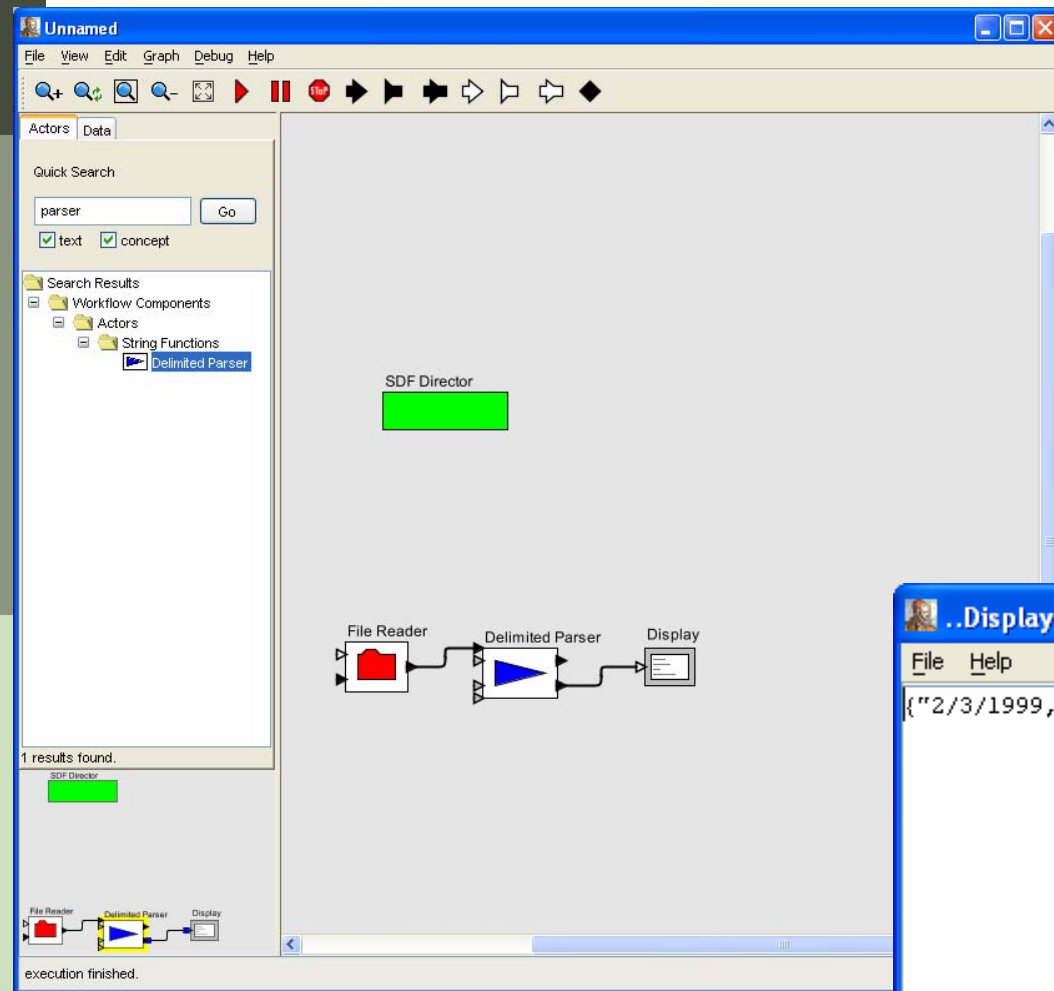
- Configure Delimited Parser output port
 - Right-click->configure ports
 - Find stringOutput output port
 - Change Type to {string} (braces are important..make sure to include them)
 - Commit
- Hook up File Reader/Delimited Parser ports
 - Connect File Reader output port to uppermost, black, Delimited Parser input port





Parse the file (cont'd)

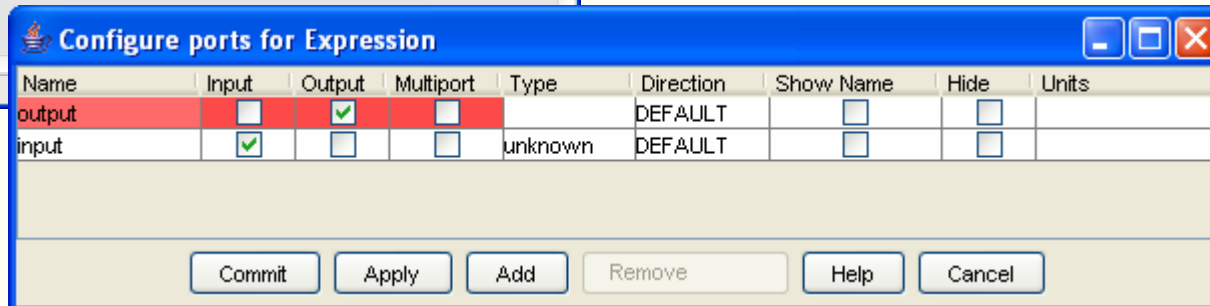
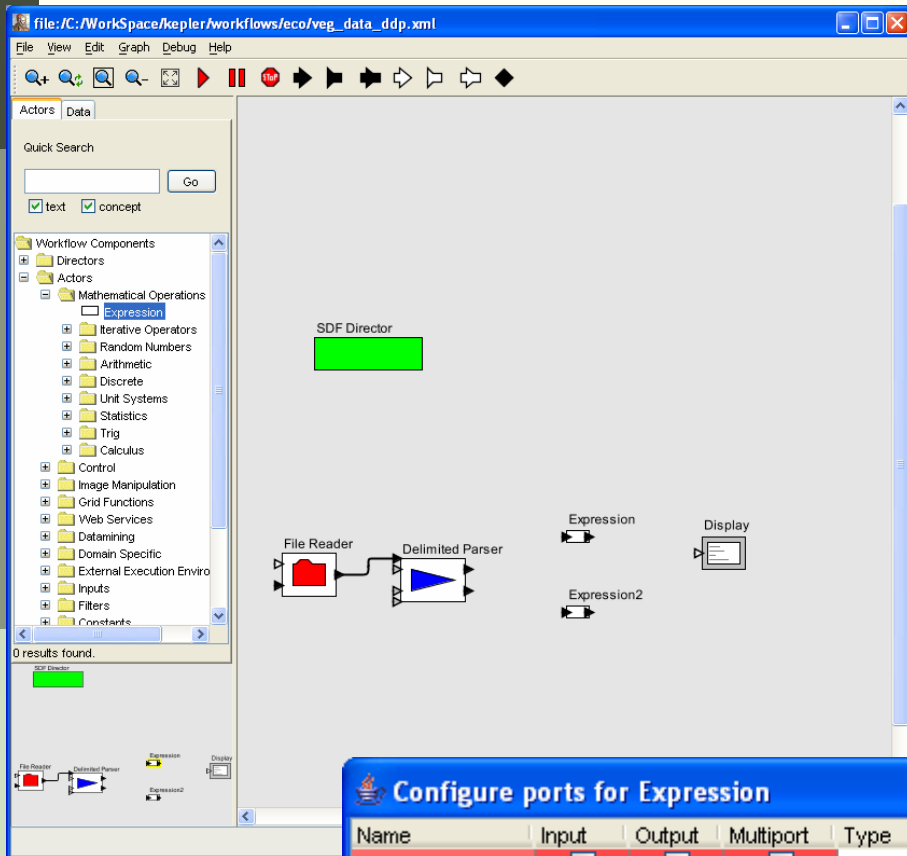
- Hook up the Delimited Parser/Display ports
Use lower, type=string output port for Delimited Parser
- Run
- Save your work!
File->Save->
browse to
c:\kepler\workflows\eco\
name as veg_data2_xxx.xml,
where xxx is your initials





Extract columns

- ❑ Delete the connection to the Display actor
- ❑ Move (but keep) Display actor
- ❑ Locate Expression actor
 - Search "expression" or Actors->Mathematical Operations
- ❑ Add Expression actor to canvas two times
 - Right-click->Configure ports->Add
 - Check Input box for new port
 - Click blank box under Name
 - Type port name "input"
 - Commit



ER





Extract columns-cont'd

□ Configure Expression actors

Double-click or right-click->configure

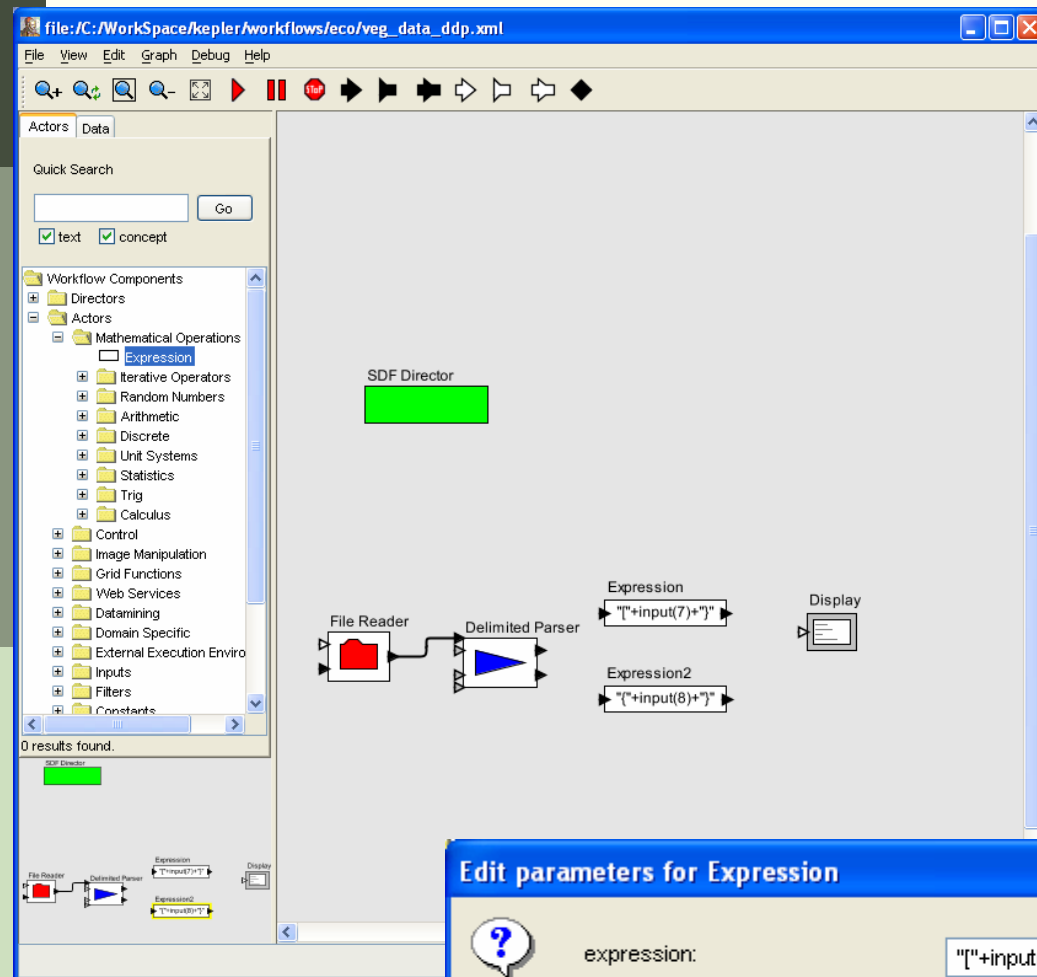
Type in first actor:

`"{" + input(7) + "}"`

Type in second actor:

`"{" + input(8) + "}"`

Commit



Edit parameters for Expression



expression:

`"{" + input(7) + "}"`

Commit

Add

Remove

Restore Defaults

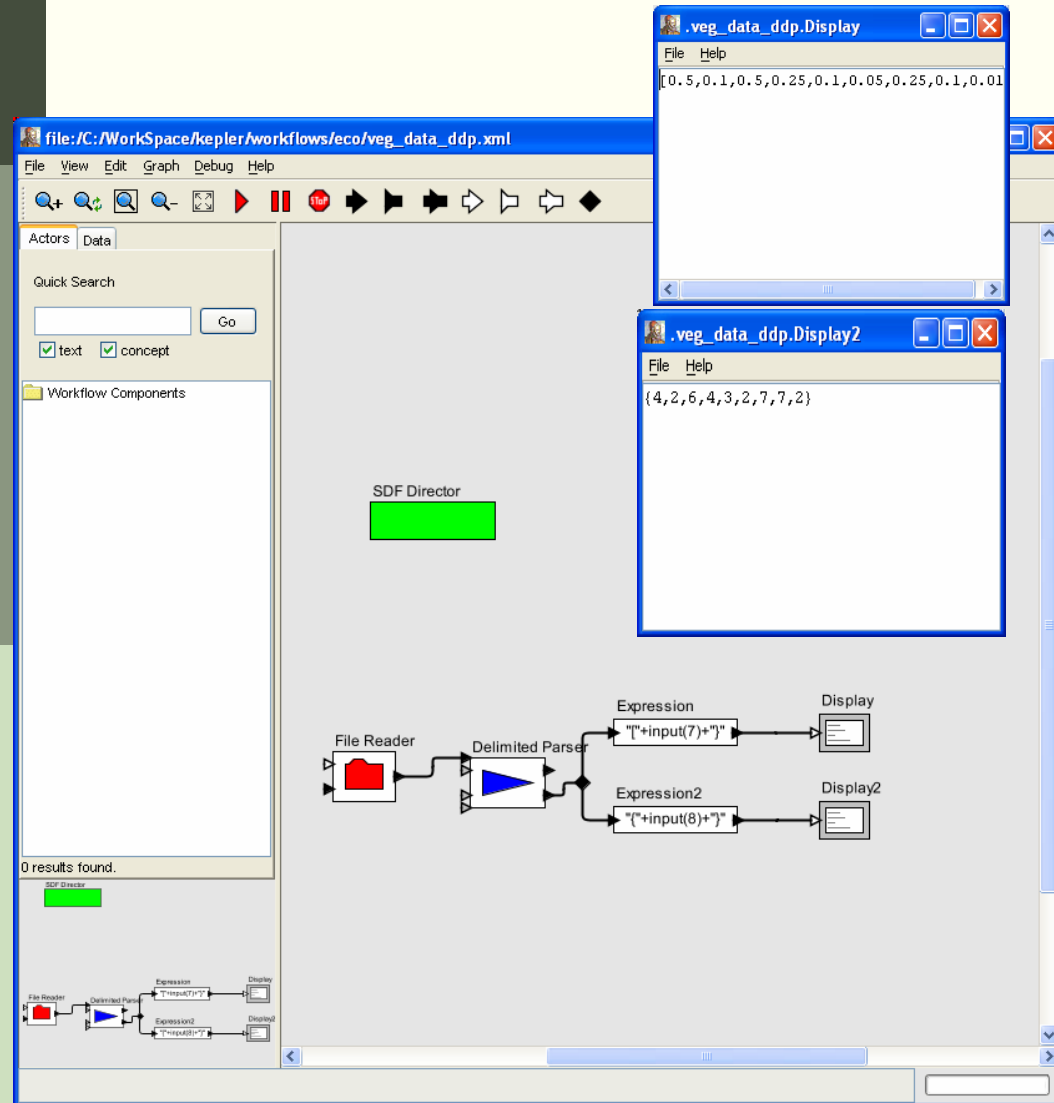
Preferences

Help

Cancel



Extract columns-cont'd



- Add relation
 - Click (one time) on diamond on task bar
 - Hook Delimited Parser lower (stringOutput) output port to relation
- Hook relation to Expression input ports
 - Hint: it's easier to click on the Expression port and drag to the relation than vice versa
- Copy the Display actor
 - Click on it
 - Copy and paste (Edit menu or shortcuts)
 - Move copy so both show
- Hook output ports on each Expression actor to the input port of a Display actor
- Run
- Move the top Display window so you can see both





Convert to decimal number

The screenshot shows the SEEK software interface. The main window displays a workflow diagram with the following components: File Reader, Delimited Parser, Expression (with text "[*input(7)+]"), Expression To Token, Expression2 (with text "[*input(8)+]"), and Expression To Token2. Two windows are open: ".veg_data_ddp.Display" showing the list {0.5, 0.1, 0.5, 0.25, 0.1, 0.05, 0.25, 0} and ".veg_data_ddp.Display2" showing the list {4, 2, 6, 4, 3, 2, 7, 7, 2}. The left sidebar shows a search for "converter" with results including "Expression To Token".

- ❑ Add Expression To Token actor two times
 - Search for "converter"
- ❑ Configure ports
 - Input Type=string
 - Output Type={ double}
 - Commit
- ❑ Hook up Expression output ports to Expression To Token input ports
- ❑ Hook up Expression To Token output ports to Display input ports
- ❑ Run

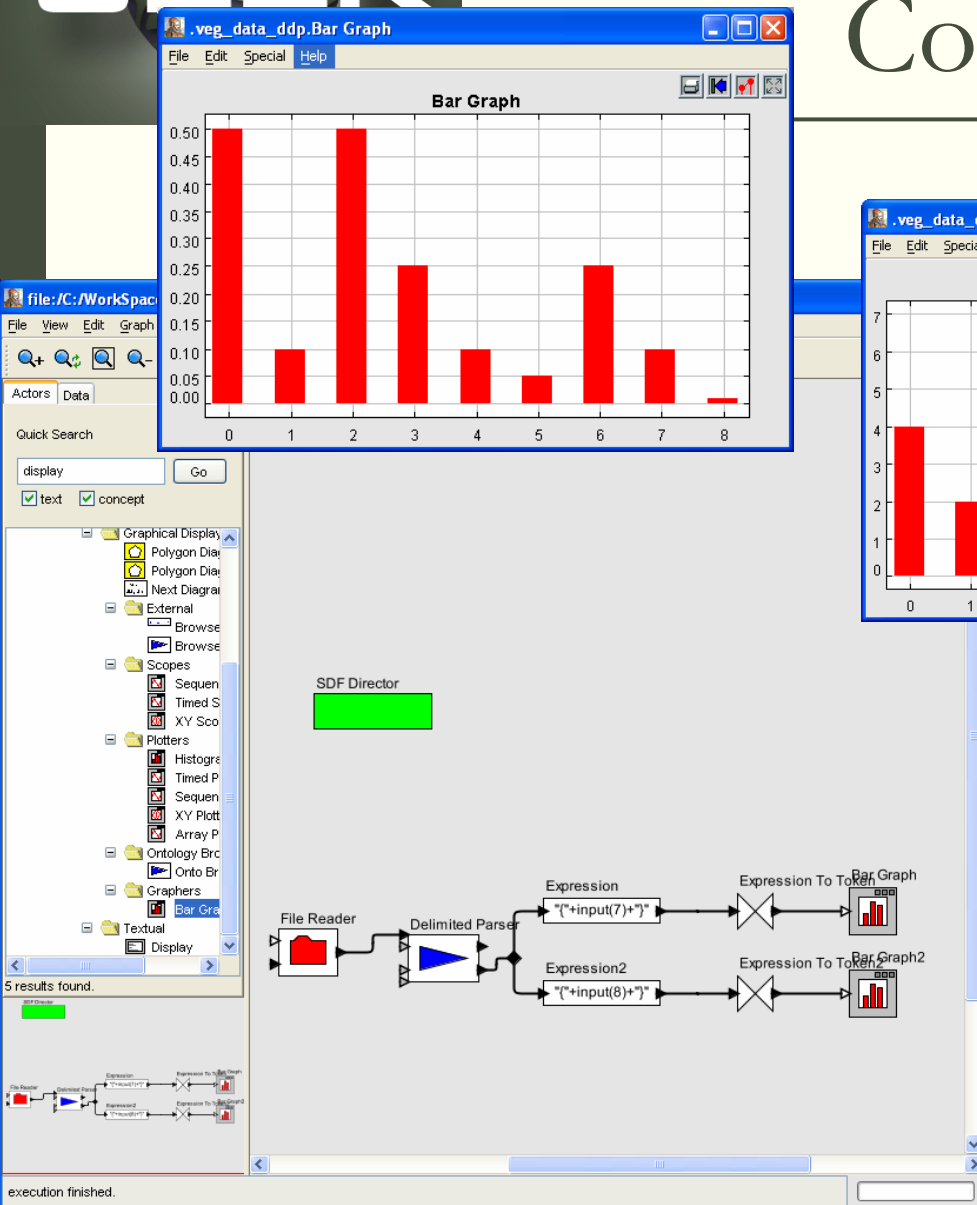
The dialog box titled "Configure ports for Expression To Token" contains a table with the following data:

Name	Input	Output	Multiport	Type	Direction	Show Name	Hide	Units
input	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	string	DEFAULT	<input type="checkbox"/>	<input type="checkbox"/>	
output	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	{double}	DEFAULT	<input type="checkbox"/>	<input type="checkbox"/>	

Buttons at the bottom: Commit, Apply, Add, Remove, Help, Cancel.



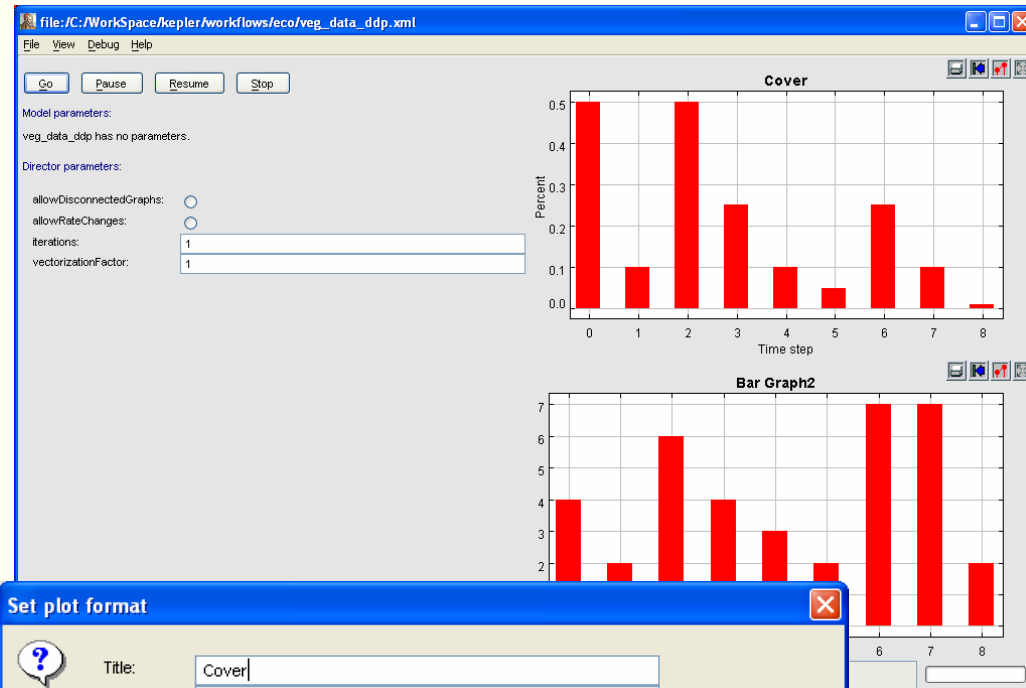
Construct bar graphs



- ❑ Delete Display actors
- ❑ Add Bar Graph display actor two times
- ❑ Hook up ports
- ❑ Run



Configure Plots



- Open Run Window
View->Run Window
- Reformat Plots
Click small icon with red dots
"Set plot format"
Change Title, label x and y axis
- Close Run Window
- Run again
- Save

The 'Set plot format' dialog box is shown with the following settings:

- Title: Cover
- X Label: Time step
- Y Label: Percent
- X Range: 0.0, 8.0
- Y Range: 0.0, 0.5
- Marks: ☒ none ☐ points ☐ dots ☐ various ☐ pixels
- X Ticks: (empty field)
- Y Ticks: (empty field)
- Grid: ☒
- Stems: ☐
- Connect: ☐
- Use Color: ☒

Buttons: Apply, Cancel





Conclusion

- ▣ Take home message....

Use EML to describe datasets...
Some pain now vs. much pain later

