



はじめての **Kepler**

Kepler 1.0.0
May 12, 2008

日本語訳
Kepler 勉強会
(大場真・小川安紀子)

January 6, 2009



はじめての Kepler

はじめての *Kepler* ガイドは、科学ワークフローを作成・実行したい科学者のためのチュートリアル形式のマニュアルです。

この日本語版は、Kepler 勉強会メンバーにより作成されました（分担は 1, 2 章、6 章: 6.3-6.5、7 章: 7.1-7.2 が小川、3,4,5 章、6 章: 6.1-6.2、7 章: 7.3-7.4、8 章が大場）。

目次

はじめての Kepler.....	1
1. はじめに.....	3
1.1. Kepler とは?	3
1.2. 科学ワークフローとは?.....	5
2. Kepler のダウンロードとインストール.....	6
2.1. システム要件.....	6
2.2. Windows にインストール	7
2.3. Macintosh にインストール.....	7
2.4. Linux にインストール	8
3. Kepler の開始.....	8
3.1. Windows と Macintosh	9
3.2. Linux.....	9
4. Kepler の基本コンポーネント.....	9
4.1. ディレクターとアクター(Director, Actor).....	10
4.2. ポート(Port).....	11
4.3. リレーション(Relation)	12
4.4. パラメーター(Parameter).....	12
5. Kepler インターフェース.....	12
5.1. ツールバー.....	13
5.2. コンポーネント・データアクセスエリア(Components and Data Access Area) 14	
5.3. ディレクター(Director)とアクター(Actor)アイコン	15
5.4. ワークフローキャンバス(Workflow Canvas).....	17
6. Kepler での基本操作.....	17

6.1. 既存のワークフローを開く.....	18
6.1.1. 例 1: Lotka-Volterra(ロトカ・ヴォルテラ) ワークフロー.....	18
6.2. 既存の科学ワークフローの実行.....	20
6.2.1. 例 2: デフォルトパラメーターでロトカ・ヴォルテラ ワークフローを 実行する.....	21
6.2.2. 例 2: 修正したパラメーターでロトカ・ヴォルテラ ワークフローを実 行する 22	
6.3. 既存の科学ワークフローを編集する.....	26
6.3.1. 例 4: Image J ワークフローの解析プロセスを編集/置き換える.....	27
6.4. Kepler で検索する.....	28
6.4.1. 利用可能なデータを検索する.....	28
6.4.2. 利用可能なプロセッシング・コンポーネントを検索する.....	30
6.5. 基本的な科学ワークフローを作成する.....	31
6.5.1. 例 5: “Hello World (ハローワールド)” ワークフローを作成する....	31
6.5.2. 例 6: 自分のデータを使って簡単なワークフローを作成する.....	33
7. サンプル科学ワークフロー.....	35
7.1. サンプルワークフロー1 – 基本統計量.....	35
7.2. サンプルワークフロー2 –線形回帰.....	37
7.3. サンプルワークフロー3-Web サービスとデータ変換.....	42
7.4. サンプルワークフロー4-Kepler からの外部アプリケーションの実行 (External Execution(外部実行)アクター).....	49
8. 付録.....	51
8.1. Ptolemy II(トレミー2)-Kepler の基礎.....	51
8.2. アクターリファレンス.....	52

1. はじめに

“はじめての Kepler”ガイドは Kepler の主なコンポーネントと機能性を紹介することを目的としており、科学ワークフローをまず使ってみて、修正し、最終的にはあなた自身のワークフローを作るための段階的なインストラクションが書いてあります。このガイドでは、アプリケーション特有の用語や概念はもちろん、アプリケーションのインターフェイスを簡単に紹介しています。とりあえず Kepler の一般的な原理に慣れることができたら、Section 7 で紹介するサンプルワークフローのいくつかを試してみることをおすすめします。ワークフローのコンポーネントを使ったり修正したりするのがどんなに簡単かを感じ、また、コンポーネントを組み合わせることによってパワフルなワークフローができあがるプロセスを体験できるでしょう。

1.1. Kepler とは？

Kepler は科学データの解析とモデリングのためのソフトウェア・アプリケーションです。Kepler はこれらの解析プロセスを視覚的に表現することにより、実行可能モデルを構築

する作業を単純化します。これらの表現、つまり、それぞれ独立した解析やモデル要素の間のデータの流を表示するものが“科学ワークフロー”です(図1)。

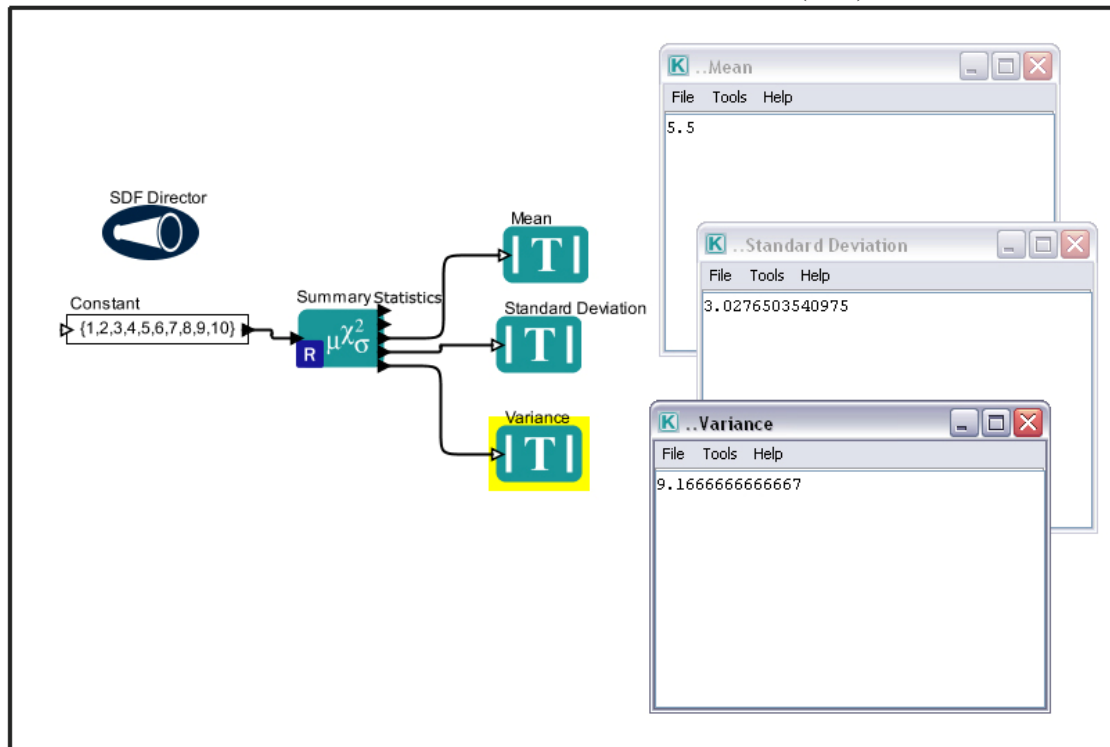


図 1: Kepler で作成された簡単な科学ワークフロー。

Kepler を使えば、科学者は単純なドラッグ・アンド・ドロップでコンポーネントをワークフロー作成エリアに持ってきてそれらを連結することで、自分の解析の視覚的なモデルを表すデータフローを組み立て、実行可能な科学ワークフローを作ることができます。Kepler はワークフロー全体を視覚的に表現できるので、データがあるコンポーネントから別なコンポーネントへどのように流れているのかを容易に確認することができます。出来上がったワークフローはテキスト形式で保存し、e-mail で研究仲間に送ったり、公開して世界中の科学者と共有することができます。

Kepler のユーザはコンピュータ科学に関する知識や経験がなくても、標準的なコンポーネントでワークフローを作ったり、既存のワークフローを自分の目的に合うように修正したりすることができます。定量的な解析では R や他の統計解析のプログラムを作成したり共有するために視覚的インターフェイスを使うことができます。ユーザは R のプログラミングの方法を知らなくてもそのパワフルな解析機能を利用することができます。すでにプログラムされた Kepler コンポーネントを、視覚的に表現されたワークフローにただドラッグするだけでよいのです。上級ユーザももちろん、Kepler の多くの利点、特に、複雑なプログラムや解析プロセスをわかりやすく共有しやすい形で提示することに関しての利点に気づくことでしょう。

Kepler は、科学者が他の科学者とデータやワークフローを共有したり、世界中の他の科学者のデータや解析ワークフローを利用したりすることを可能にする分散型コンピューティング技術を使っています。Kepler はまた、絶えず拡大しつづける地理的に分散データ・リポジトリ（集積所）やコンピュータ・リソース、そしてワークフロー・ライブラ

り（たとえば、フィールド施設で収集された生態学データや博物館の標本コレクション、または地球科学データなど）へのアクセスを提供します。

1.2. 科学ワークフローとは？

科学ワークフローとは、科学データ（リアルタイム・センサー・データ、医学や衛星画像、シミュレーション出力、観測データなど）にアクセスしたり、取り込んだデータに対して複雑な解析を実行したりするための便利なツールです。

それぞれのワークフローは、データベースへのアクセス、検索、データ解析とマイニング、そして高性能コンピュータ群で行う集中的な演算などの解析ステップで構成されています。それぞれのワークフロー・ステップは、“アクター”と呼ばれる処理コンポーネントで表され、Keplerの視覚的なインターフェイス上でドラッグ・アンド・ドロップによりワークフローに組み込むことができます。科学者は、アクター（または別のセクションで後述する他のコンポーネント）が連結されてできたワークフローを使って、計算が実行されるのと同時に、それを検査したり表示したりしながら、必要ならパラメータを調整し、(数値)実験を再実行して結果を表示することができます。¹

ワークフローは理論的なモデルを表現したものである場合もあれば、観測データの解析プロセスである場合もあります。ワークフローは単純で線形である場合も、複雑で非線形である場合もあります。科学ワークフローの利点の一つに、入れ子構造にすることが可能という点があります。つまり、いうなれば“サブワークフロー”を格納し、組み込まれたタスクを実行させることができるのです。入れ子ワークフロー（コンポジットアクターとも言う）は、複雑なタスクもこなす能力を秘めた再利用可能なコンポーネントです。

Keplerの科学ワークフローは、（分散するデータやコンピュータサービスなどの資源へのアクセスを提供することにより）今日のGrid技術の恩恵を得る手段を提供しつつ、しかもその裏にあるこれらの技術の複雑さを表から見えないようにしています。ケプラーは、科学者が興味のある科学的質問に集中することができるよう、低レベルのデータ処理を自動化してくれるのです。

ワークフローには以下のような長所もあります。

- 一つの解析をあらゆる側面から文書化
- 解析ステップの視覚的表現
- 複数のシステム上で駆動する能力
- 最小限の努力であるプロジェクトを再現する能力
- 異なるプロジェクトで既存のワークフローの一部もしくは全部が再利用可能

今日、ほとんどの科学ワークフローは、多様なソフトウェアプログラムと高度なプログラミング言語がないと使うことができません。これまでの研究者は図式的にシステムをモデル化するためにはSTELLAやSimulinkなどを、また、統計解析を実行するためにRやMATLABなどを使ってきました。ユーザによっては計算にExcelなどを使っています。Excelはユーザフレンドリーですが、どのようなステップでその計算が実行されたかの記録を残してはくれません。Keplerはこれらすべてのプログラムの利点を合体させ、

¹ See Ludäscher, B., I. Altintas, C. Berkley, D. Higgins, E. Jaeger-Frank, M. Jones, E. Lee, J. Tao, Y. Zhao. 2005. Scientific Workflow Management and the Kepler System, DOI: 10.1002/cpe.994

ユーザがモデリングや解析を行い、データを表示させるといった作業を使いやすいインターフェイスで行うことを可能にします。

Kepler は Ptolemy II (トレミー) というオープン・ソースの視覚的モデリングシステム (<http://ptolemy.eecs.berkeley.edu/ptolemyII/>) をベースに構築されているため、科学者に単一の作業環境を提供します。この結果、いくつかの異なるソフトウェア・プログラムを統合したり、コンピュータ・プログラマーの助けを必要とせずにワークフローを構築したりすることができるユーザ・フレンドリーなプログラムが出来あがったのです。

Kepler には、一般的な算術、統計、信号処理、データ・インプット、データ操作、表示など、すぐに使えるコンポーネントが数多く標準装備されています。R や MATLAB ベースの統計解析、画像処理、また、GIS 機能も外部パッケージへの直接リンクを通じて利用可能です。また、まったく新しいコンポーネントを作ったり、他のプログラム (たとえば C 言語プログラムなど) のコンポーネントを Kepler で使えるように組み込むことも可能です。

2. Kepler のダウンロードとインストール

Kepler は、オープンソースの、しかも OS が Windows でも、Macintosh でも、Linux でも使える、OS に依存しないソフトウェア・プログラムです。Kepler は以下のウェブサイトからダウンロードできます。 <http://kepler-project.org> . Kepler 1.0.0 が最新版リリースです。

Kepler のリリースは常に継続して行われている作業の成果です。Kepler のユーザも便利な新しい機能を提案したり、バグやその他の問題をデザイナーに報告したりすることで開発に参加し、この成果に貢献することができます。詳しい情報は <http://www.kepler-project.org/Wiki.jsp?page=GettingInvolved> をご覧ください。Kepler 開発にユーザコミュニティが関わることは、システムを実際に利用している科学者のニーズに適応させていくために、とても価値のあることです。変更や更新を見逃さないためには、以下のリンクから Kepler user メーリングリストを購読してください。

<http://mercury.nceas.ucsb.edu/ecoinformatics/mailman/listinfo/kepler-users> .

2.1. システム要件

Kepler を実行するのに推奨されるシステム要件は以下のとおりです:

- 300 MB のハードディスク・スペース
- 最低 512 MB の RAM、1 GB 以上を推奨
- 最低 2 GHz の CPU
- Java 1.5.x 以降 (Java 1.6 は使わないで下さい)
- ネットワーク接続環境 (オプション). Kepler を実行するのにネットワーク接続は必要ありませんが、ワークフローの多くがネットワーク・リソースにアクセスする必要があります。)
- R ソフトウェア (オプション)。R は統計計算とグラフ作成のためのプログラム言語とそれを使うための環境で、Kepler のよく使われるいくつかの機能が必要とします。R はウィンドウズとマック用の Kepler フルインストーラに含まれていますが、Linux 用インストーラには含まれていません。

Kepler をダウンロードしてインストールするには、以下の中から、あなたのシステムに適した手順に従ってください。あなたの接続環境によってはインストーラのダウンロードに時間がかかるかもしれません。

注意： Java 1.5.x が必要です。これは以下の Sun の Java ウェブサイトから入手可能です <http://java.sun.com/j2se/downloads/>。もしくはあなた組織のシステム管理者から入手してください。ウィンドウズと Linux の Kepler インストーラは、もし Java がまだあなたのシステムにインストールされていなかったら、自動的に Java 1.5.x のダウンロードページを開きます。詳しい情報を得るにはあなたの OS のインストールマニュアルをご覧ください。

2.2. Windows にインストール

Windows 用インストーラは、Kepler と **R** (オプション) -- いくつかの Kepler のアクターに使われている統計計算言語・環境 -- をあなたのシステムにインストールします。もしあなたがまだ Java 1.5.x をインストールしていなかったら、インストーラはこれをダウンロードするためのページを開きます。Java 1.5.x は Kepler ソフトウェアを実行するために必要です。

もし R が Kepler と一緒にインストールされた場合、既にインストールしてある R のバージョンに影響はないはずですが、ただし、Kepler インストーラは、R をコマンドラインから ('R' と入力することにより) 起動する時には、新しいバージョンの R が起動されるようにあなたのシステムを更新します。既に設置してあるショートカットは先にインストールされた R アプリケーションを開きます。Kepler インストーラに含まれている R のバージョンは 2.6.2 です。

次のステップに従って Windows 版 Kepler をダウンロードし、インストールしてください。

1. 次のリンクをクリック: <http://kepler-project.org/Wiki.jsp?page=Downloads> し、ウィンドウズ用インストーラを選択する
2. インストールファイルを自分のコンピュータに保存する
3. インストールファイルをダブルクリックしてウィザードをインストールする
4. 表示される指示に従って Kepler のインストールを完了する
5. インストールプロセスが完了したら、Kepler のショートカットアイコンがデスクトップとスタートメニュー (の両方か、どちらか) に現れます(図2)。



図 2: Kepler ショートカットアイコン

2.3. Macintosh にインストール

Macintosh 用インストーラは、Kepler と **R** (オプション) -- いくつかの Kepler のアクターに使われている統計ソフト -- をあなたのシステムにインストールします。Java はすでに Mac OS X のオペレーティング・システムの一部に含まれていますので、インストールの必要はありません。

R はよく使われるいくつかの Kepler の機能に必要とされるため、Kepler のインストールと同時にインストールすること（初期設定）をおすすめします。もし R が Kepler と一緒にインストールされた場合、既にインストールしてある R のバージョンに影響はないはずです。ただし、Kepler インストーラは、R をコマンドラインから（'R' と入力することにより）起動する時には新しいバージョンの R が起動されるようにあなたのシステムを更新します。既に設置してあるショートカットは先にインストールされた R アプリケーションを開きます。Kepler インストーラに含まれている R のバージョンは 2.6.2 です。次のステップに従って Macintosh 版 Kepler をダウンロードし、インストールしてください。

1. 次のリンクをクリック：<http://kepler-project.org/Wiki.jsp?page=Downloads> し、Mac 用インストーラを選択する
2. 解凍の完了と共にデスクトップに現れるインストールアイコンをダブルクリック
3. インストールウィザードに表示されるステップに従い、Kepler のインストールプロセスを完了する

Kepler アイコンはアプリケーション/Kepler に現れます。アイコンをデスクトップやドックにドラッグ・アンド・ドロップしてもかまいません。

2.4. Linux にインストール

Linux 用インストーラは Kepler をダウンロードし、もしまだ Java 1.5.x をインストールされていないならば、これをダウンロードするためのページを開きます。Java 1.5.x は Kepler ソフトウェアを実行するために必要です。

統計計算とグラフ作成のためのソフトである R は、Linux 用インストーラには**含まれていません**。R は、いくつかのよく使われる機能に必要なため、R をダウンロードしてインストールすることを推奨します。R についての詳しい情報を得るには次のサイトをご覧ください。<http://www.r-project.org>

次のステップに従って Linux 版 Kepler をダウンロードし、インストールしてください。

1. 次のリンクをクリック：<http://kepler-project.org/Wiki.jsp?page=Downloads> し、Linux 用インストーラを選択する
2. インストールファイルを自分のコンピュータに保存する
3. インストールファイルをダブルクリックする。インストールを続ける前に他の全てのプログラムを終了することを推奨します
4. Kepler インストーラがステータスバーにインストールの進行状況を表示します。Kepler が以前にインストールされていた場合は、インストーラは全てのキャッシュファイルを上書きします

3. Kepler の開始

Keplerを開始するには、以下の中からプラットフォーム(OS)にあわせたガイドに従って下さい。

3.1. Windows と Macintosh

PCでKeplerを開始するには、デスクトップ上のKeplerのショートカットアイコンをダブルクリックして下さい(図2)。Keplerはスタートメニューからも開始できます。

Windowsなら、「スタート」ボタンを押して、「すべてのプログラム」から「Kepler」フォルダを選び、「Kepler 1.0.0」*訳注を選ぶことで開始できます。Macintoshの場合は「アプリケーション」の「Kepler」にKeplerのアイコンがあります。このアイコンはデスクトップやドックにドラッグ・アンド・ドロップすることができます。

Keplerのメイン画面が開きます(図3)。このウィンドウからサンプルや既存の科学ワークフローにアクセス・実行ができ、また独自のワークフローを作成できます。既存のワークフローを開いたり新たにワークフローを作成したりすると、常に新しいウィンドウが開きます。複数開いたウィンドウは、いくつかのワークフローで同時に作業したり、ワークフロー間で比較、コピー、ペーストしたりすることを可能にします。

*訳注 ver. 1.0.0の場合

3.2. Linux

LinuxでKeplerを開始するには、次のステップを踏みます:

1. シェルウィンドウを開きます。いくつかのLinuxシステムでは、デスクトップ上を右クリックし、「ターミナルを開く」を選ぶことで、シェルを開くことができます。もし使用しているシステムについて知りたいことがあれば、システム管理者に聞いて下さい。
2. Keplerをインストールしたディレクトリへ移動して下さい。ディレクトリを移動するにはcdコマンドを使います。(例えばcd <移動したいディレクトリ名>)
3. ./Kepler.shと入力することで、アプリケーションが開始します。

Keplerのメイン画面が開きます(図2,3)。このウィンドウからサンプルや既存の科学ワークフローにアクセス・実行ができ、また独自のワークフローを作成できます。既存のワークフローを開いたり新たにワークフローを作成したりすると、常に新しいウィンドウが開きます。複数開いたウィンドウは、いくつかのワークフローで同時に作業したり、ワークフロー間で比較、コピー、ペーストしたりすることを可能にします。

4. Keplerの基本コンポーネント

ワークフローはカスタマイズ可能な部品であるコンポーネントーディレクター(Director)、アクター(Actor)、パラメーター(Parameter)ーと、コンポーネント間を連絡するためのリレーション(Relation)、ポート(Port)から構成されます。

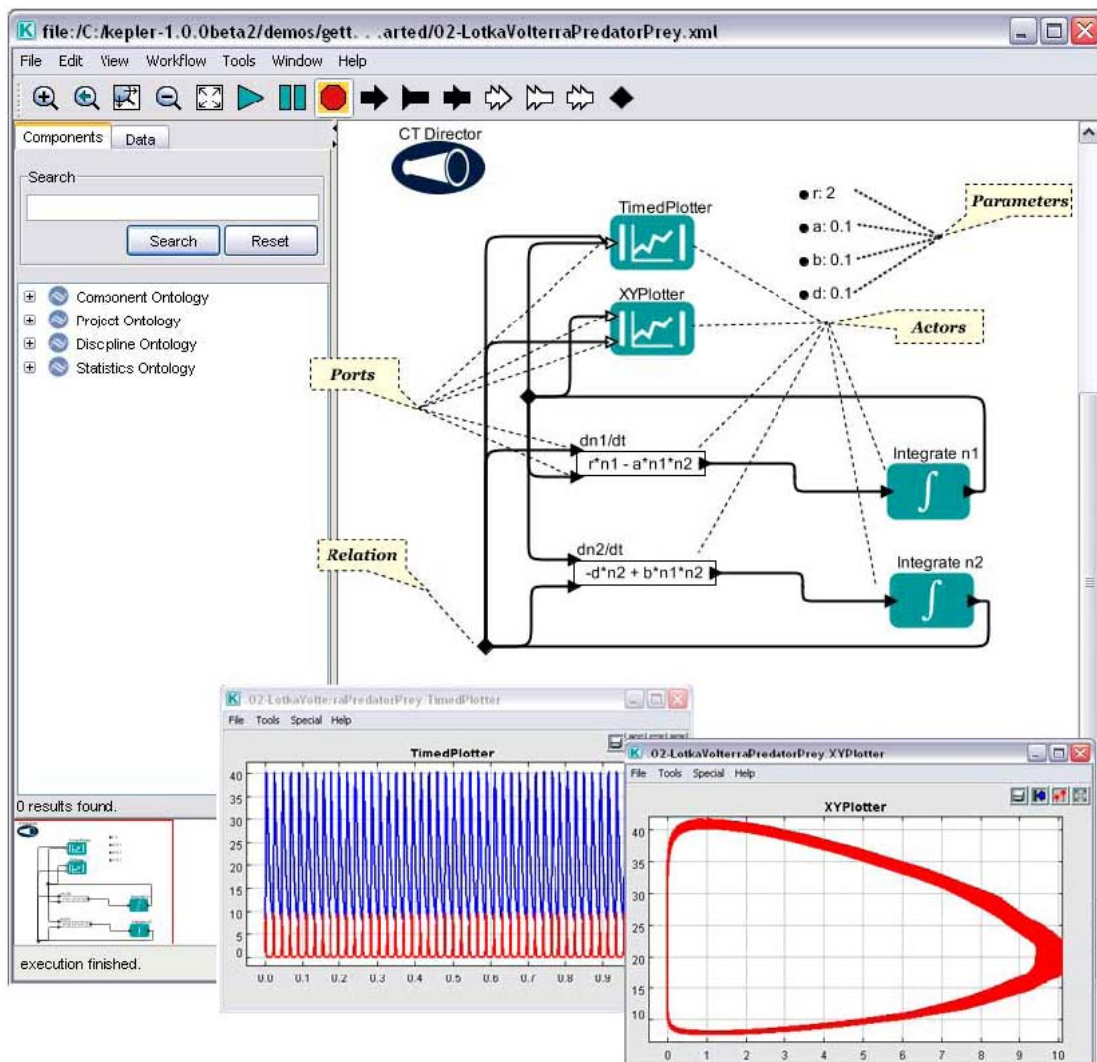


図 3: いくつかの主要なワークフローコンポーネントと Kepler の Main Window(メインウィンドウ)のハイライト。右下のウィンドウは Output Window(出力ウィンドウ)で、結果グラフを表示するためにワークフローにより生成されたものです

4.1. ディレクターとアクター(Director, Actor)

Kepler はワークフローの様々なコンポーネントを視覚的に表現するため、「ディレクター(Director, 監督)とアクター(Actor, 俳優)」という比喻を用いています。ディレクターは、映画監督が出演者とスタッフを監督するように、ワークフローの実行を制御(監督)します。アクターはディレクターより実行指令を受け取ります。つまり、アクターはどのようなプロセスが生じるのかを示し、ディレクターはそれがいつ生じるのかを特定します。すべてのワークフローは、特定の計算方法のモデルを用いるワークフローの実行を制御する適切なディレクターを持たなければなりません。Kepler の計算方法のモデルは、ディレクターによって表されます。例えば、実行を同期させたいワークフローでは、その前に計算される手順が終わったと同時に次のコンポーネントのプロセスが開始されるようなディレクターを用います(SDF ディレクター)。これ以外には例えば、ワークフロー

コンポーネントを並列的に実行させたいなら、一つ以上のコンポーネントを同時に実行できるディレクターを用います(*PN*ディレクター)。通常使用されるディレクターを Kepler はパッケージとして持ちますが、必要に応じてアクセス可能な基盤ソフト Ptolemy II にあるものも利用可能です。計算のワークフローモデルについてより詳細な議論について、「Kepler ユーザーマニュアル」か Ptolemy II のドキュメントを参照して下さい。

Composit(コンポジット)アクターは、より複雑な操作を行うためにバンドルされたアクターのコレクションあるいは集合です。コンポジットアクターはワークフローの中で、基本的にはネストしたワークフローかサブサブワークフローとして使用可能である(図4)。一つのワークフロー全体は、コンポジットアクターとして表現することができ、またカプセル化されたワークフローであるコンポジットコンポーネントとして含まれることもできます。さらに複雑なワークフローでは、コンポジットアクターやワークフローごとに異なるディレクターを持つこともできます。

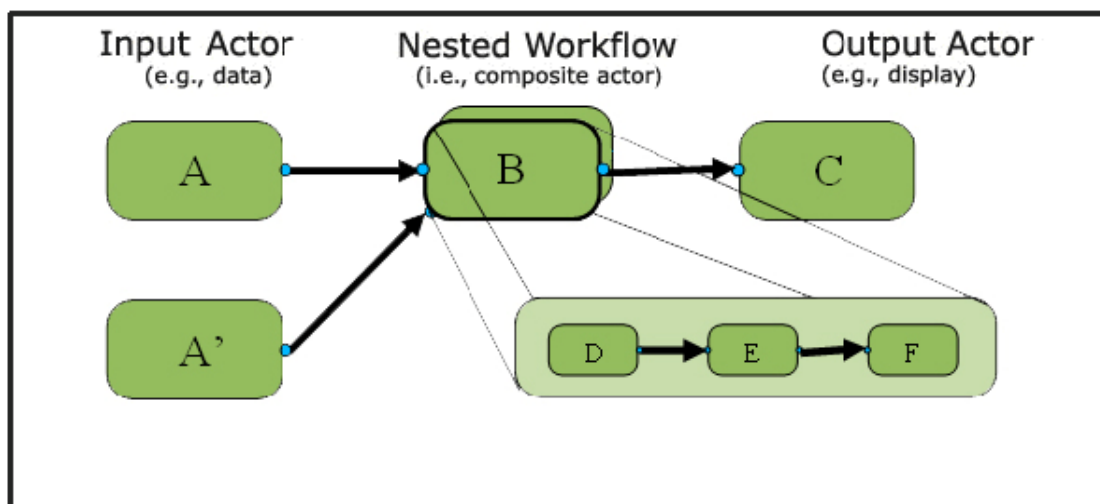


図 4: ネストしたワークフローの図解

Kepler は科学ワークフローを生成したり編集するための多くのアクターを提供します。アクターを個人的な目的で利用するために Kepler に追加したり、他の人が利用可能なようにしたりすることができます。

4.2. ポート (Port)

ワークフローのそれぞれのアクターは、データを消費したり作成したり、同じワークフローの他のアクターと通信するため、1つ以上のポートを持ちます。ワークフロー中でアクターはポートを経由して結びつきます。あるアクターのポートと他のアクターのポートとの間のデータフローに相当するリンクは、チャンネルと呼ばれます。ポートは3種類に分けられます:

- Input port(入力ポート) このアクターによるデータ消費のため;
- Output port(出力ポート) このアクターによるデータ作成ため;そして

- Input/Output port(入出力ポート) アクターによるデータの消費と作成両方のため。

どのポートも単独(singular)あるいは複数(multiple)のどちらかのポートとして設定されます。単独入力ポートは単独チャンネルのみに接続が可能で、複数入力チャンネルは複数チャンネルに接続できます。Kepler では、単独ポートは黒い三角、複数ポートは白い三角として示されます。

ワークフローは外部(external)ポートとポートパラメータも使うことができます。より詳しくは Ptolemy のドキュメントを参照して下さい。

4.3. リレーション(Relation)

リレーションはデータフローの分岐を可能にします。分岐されたデータはワークフロー内の複数の場所へ送ることが可能です。例えば、ある操作アクターの出力を次の処理のため別の操作アクターへ直接送り、かつ、そこでこのデータを表示させるために表示アクターへも送ることができます。出力用のデータチャンネルの中にリレーションをおくと、同時に2つの場所へ情報を送ることが可能です。

4.4. パラメーター(Parameter)

パラメーターは設定可能な値で、ワークフロー、ディレクターやアクターに付属させることができます。例えば *Integrator*(積分)アクターは *InitialState* と呼ばれるパラメーターを持ち、これは積分される関数の初期値を設定します。シミュレーションモデルのパラメーターアクターは、シミュレーションの状態を制御する、例えば初期値のようなパラメーターを設定することができます。ディレクターのパラメーターはワークフローの反復回数や反復を続ける条件を定めます。

次のセクションではインターフェースの概観と、様々なワークフローを開いたり、編集したり、実行したりする例を紹介します。

5. Kepler インターフェース

科学ワークフローを、Kepler 上の操作が容易なドラッグ・アンド・ドロップインターフェース上で編集したり・作成したりすることができます。Kepler アプリケーションウィンドウ(図5)の主な部分は:

- Menu bar(メニューバー) Kepler の全ての機能にアクセスすることができます。
- Toolbar(ツールバー) 最もよく利用される Kepler の機能にアクセスすることができます。
- Components and Data Access area(コンポーネント・データアクセスエリア)
Components(コンポーネント)タブと Data(データ)タブからなります。どちらのタブも機能の検索や利用可能なコンポーネントのライブラリの表示、その検索結果を表示します。

- Workflow canvas(ワークフローキャンバス) ワークフローを表示したり作ったりする場所です。
- Navigation area(ナビゲーションエリア) ワークフロー全体を表示します。ここに表示されているワークフローの部分をクリックすると、ワークフローキャンバスでその部分が選択され表示されます。

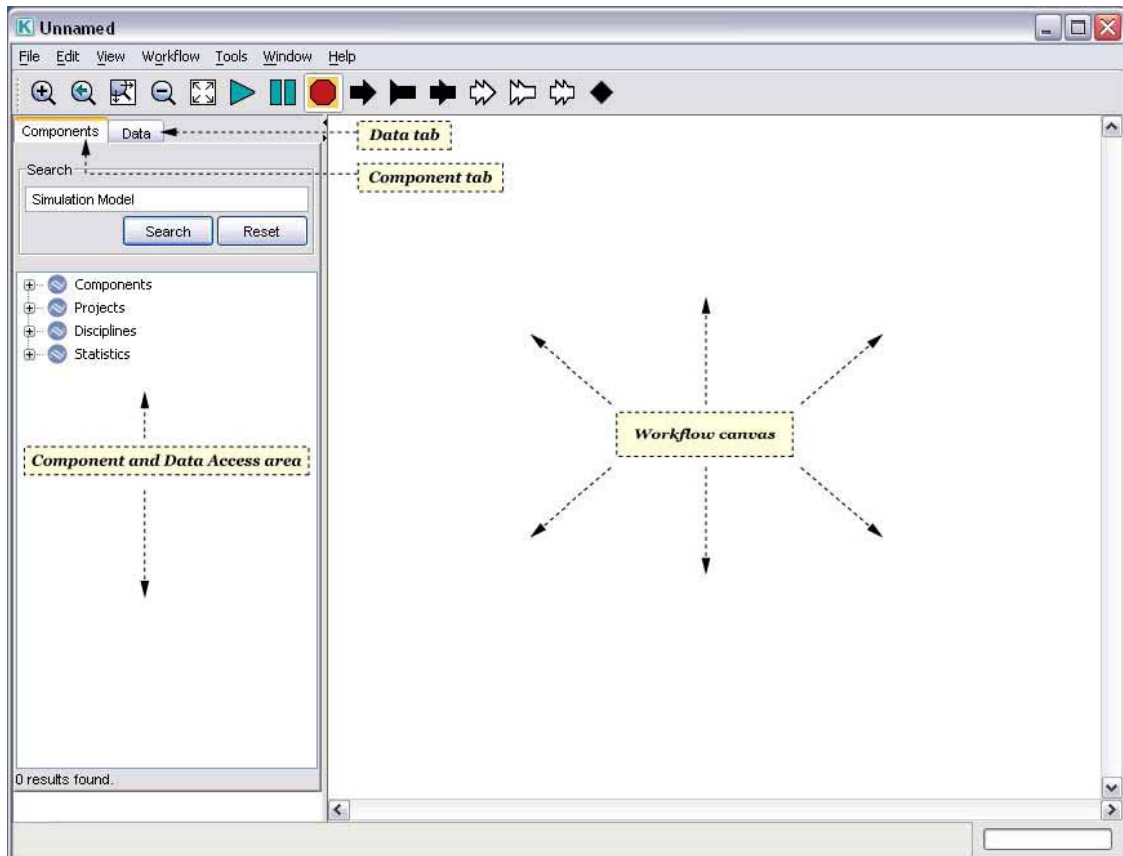


図 5: 空の Kepler ウィンドウ。主な領域を図示した。

5.1. ツールバー

ツールバーにはよく使われる Kepler の機能が入っています(図6)。

ツールバーの主な部分は:

- Viewing(表示) ワークフローキャンバスでのワークフローの Zoom In(ズームイン)、Zoom Reset(ズームリセット)、Zoom Fit(ズームフィット)、Zoom Out(ズームアウト)
- Run(実行) Run Window(実行ウィンドウ)を開かずに、ワークフローの Run(実行)、Pause(一時停止)、Stop(停止)
- Ports(ポート) ワークフローへ singular(黒三角)か multiple(白三角)の入出力ポートの追加。ワークフローへのリレーションの追加

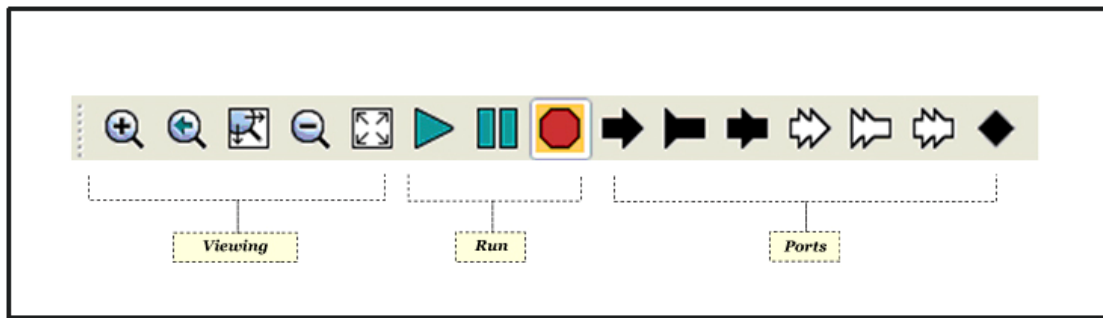


図 6: Kepler ツールバーの説明

5. 2. コンポーネント・データアクセスエリア (Components and Data Access Area)

コンポーネント・データアクセスエリアはワークフロー コンポーネントのライブラリ (「コンポーネント」タブのなかのディレクターやアクターなど)と、データセットを見つけて利用するための検索機能(「データ」タブ)があります。アプリケーションを最初に開いたとき、「コンポーネント」タブが表示されます。

Kepler のコンポーネントは 4 つの高次カテゴリーに整理されています: コンポーネント (機能ごと)、プロジェクト、専門、統計(表 1)。どのコンポーネントも複数のカテゴリーに分類され、コンポーネントツリーの複数の場所に現れる場合があります。同じコンポーネントは、複数のカテゴリーにあったとしても、どれを使っても同じこととなります。

コンポーネントツリーの中をクリックすることでコンポーネントを閲覧したり、特定のコンポーネントを見つけるために「コンポーネント」タブ上部の検索機能を使ったりすることができます。コンポーネントの探し方についての詳細は、セクションの 6.4.2 を参照して下さい。

カテゴリー (Category)	説明
コンポーネント (Components)	全コンポーネントの基本的なライブラリであり、機能によってファミリーに分類されています
プロジェクト (Projects)	プロジェクトに特有のコンポーネントのライブラリがあります (e.g., SEEK or CIPRes)
専門 (Disciplines)	コンポーネントライブラリが専門分野(例えば化学や生態学)によって分類されたものがあります
統計 (Statistics)	統計解析のためのコンポーネントライブラリがあります






表 1: Kepler のコンポーネントカテゴリー





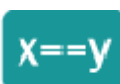


「データ」タブをクリックすると Data Access area(データアクセスエリア)が表示されます。リモートホストにあるデータセットを取り込むために EarthGrid を簡単に検索できます。データ検索の詳細はセクション 6.4.1 を参照してください。

5.3. ディレクター(Director)とアクター(Actor)アイコン

Kepler では、アイコンがコンポーネントの機能を視覚的に表現しています。ディレクターはただ一種類のアイコンで表現されます。アクターは機能別カテゴリすなわちファミリーという分類によって分けられ、それぞれのファミリーに視覚的に関連付けられたアイコンが割り振られています(表2)。

いくつかのアクターのファミリーはそのファミリーに共通のシンボルを持ちますが、独自のシンボルがないファミリーもあります。アクターアイコンの大部分は青緑色の長方形を使います。データ・ファイルアクセスアイコンのような一部のアイコンは違う色・違う形を使います。以下の表では、ファミリー・シンボルがあるものにはそう示しました。ファミリー・シンボルを持たないファミリーには、そのファミリーでのアイコンの 1 サンプルを示しました。表にあるすべてのファミリーのすべてのアイコンは、Kepler ユーザーマニュアルの 5 章で見つけることができます。

アイコン	ファミリー名	説明
	ディレクター (Director)	スタンドアローン(単独)のコンポーネントで、他のコンポーネント(アクター)に実行を指示します
	配列 (Array)	配列アクターは波カッコ {} で示されます。このファミリーに入るアクターの中には一般的な配列処理のために使われます(例えば、配列のソート)
	コンポジット (Composite)	コンポジットアクターは、複数のアクターを意味するために、複数の青緑色の長方形で表現します。コンポジットアクターは、カプセル化されたワークフローの中で、より複雑な操作(例えばサブワークフロー)を実現するためにアクターをまとめたコレクションです
	制御(Control)	制御アクターは永続的なファミリーシンボルを持ちません。これらのアクターはワークフローを制御するために使われます(例えば、停止、一時停止、反復)
	データ・ファイルアクセス (Data/File Access)	データ・ファイルアクセスアクターは永続的なファミリーシンボルを持ちません。このファミリーに入るアクターは、データの読み込み、書き込み、クエリ(問い合わせの要求)です。ここで示されているアイコンはデータ書き込みアイコンです

	データ処理 (Data Processing)	データ処理アクターはデータの合成(アセンブル)、 分解(ディスアセンブル)、更新です
	表示 (Display)	表示アクターは垂直線 2 本で示されます。このファミリーに入るアクターはテキストやグラフでワークフローの結果を出力します
	ファイル管理 (File Management)	ファイル管理アクターは永続的なファミリーシンボルを持ちません。このファミリーに入るアクターは、例えばファイルの位置を探したり zip 圧縮を解凍したりします。ここで示したアイコンはディレクトリを一覧するアイコンです
	GAMESS	GAMESS アクターはコンピューター化学ワークフローに使われます
	一般 (General)	他の分類に入らないアクターは一般ファミリーに入ることになります。一般アクターは、例えば、e-mail、ファイル操作、変換アクターです。ここで示したアイコンはフィルターアイコンです
	GIS/空間解析 (GIS/Spatial)	GIS/空間解析アクターは地理空間情報の処理に使われます
	画像処理 (Image Processing)	画像処理アイコンは画像ファイルの操作に使います
	論理 (Logic)	論理アクターは永続的なファミリーシンボルを持ちません。このファミリーに入るアクターはブール代数スイッチや論理関数です。ここで示したアイコンは等号アイコンです
	数学 (Math)	数学アイコンは永続的なファミリーシンボルを持ちません。このファミリーに入るアクターは加算、減算、積分、統計関数です。ここで示したアイコンは統計関数をあらわすために使われます(例えば、 <i>Quantizer</i> (量子化)アクター)
	モデル (Model)	モデルアクターは実線矢印を使います。モデルアクターは、統計モデル、数理モデル、ルールベースモデル、確率モデルを含みます。このアイコンは他のシンボルを加えることで、アクター機能をより識別しやすくします


	分子処理 (Molecular Processing)	分子処理アクターは左上隅の分子のマークによって示されます
	その他・外部プログラム (Other/External Program)	その他プログラム(例えば Kepler に組み込んだ C 言語プログラム)と外部プログラムは、紫色の長方形マークによって示されます。外部プログラムアクターは R, SAS, MATLAB アクターなどがあります。ここで示したアイコンは R アイコンです
	文字列 (String)	文字列アイコンは「string()」というテキストによって示され、様々な方法で文字列を操作するために使われます
	ユーティリティ (Utility)	ユーティリティアクターはレンチのマークで示されます。ユーティリティアクターはアプリケーションの特定の状況を管理したり調節したりすることを補助します
	Web サービス (Web Services)	Web サービスアクターはワイヤーフレーム地球のマークで示されます。このファミリーのアクターはリモートサービスを実行します
	単位 (Units)	単位コンポーネントは単位系を定義します

表 2: 主な Kepler のアイコン

5.4. ワークフローキャンバス (Workflow Canvas)

ワークフローキャンバスでは科学ワークフローを開いたり、作成したり、修正したりすることができます。コンポーネントは、コンポーネントエリアやデータアクセスエリアから簡単にドラッグしてキャンバスの置きたい場所へドロップすることができます。どのコンポーネントもアイコンによって表示され(例はセクション 5.3)、コンポーネントを単純に識別することができます。コンポーネント間の接続(つまりチャンネル)も視覚的に表示されるためデータと処理のフローが明確になります。

既存のワークフローを開いたり、あるいは新しいワークフローを作成したりする時には、新しいアプリケーションウィンドウが開きます。複数ウィンドウは複数のワークフローでの同時作業を可能にし、ワークフローキャンバス間でのコンポーネントの比較、コピー、貼り付けもできます。

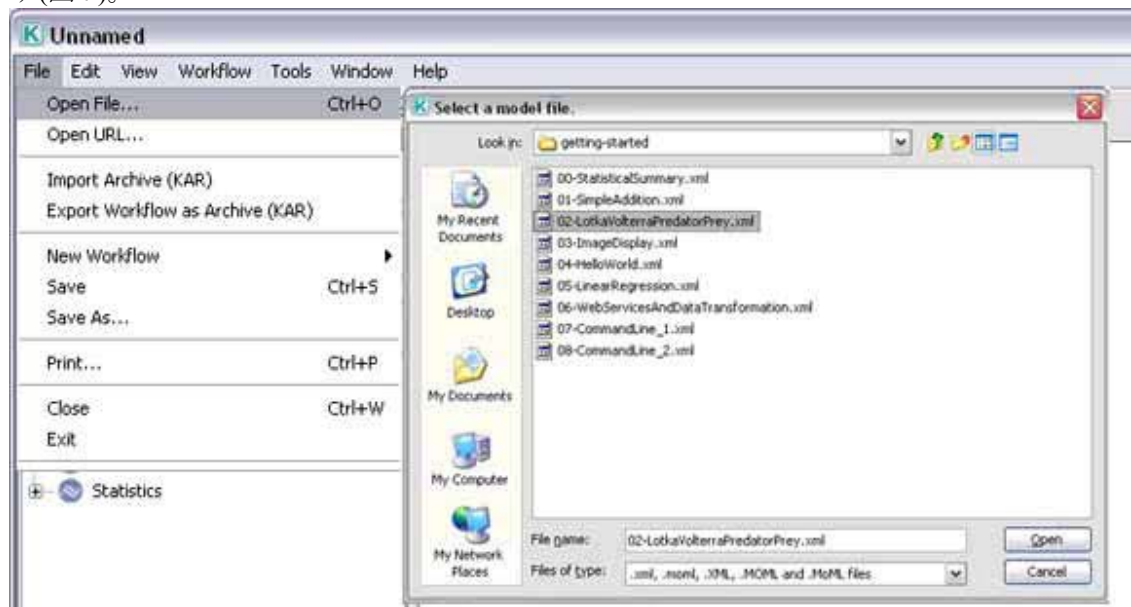
6. Kepler での基本操作

このセクションでは Kepler での基本操作、例えば既存のワークフローを開いたり、実行させたり、あなた自身のワークフローを編集したり、設計したり、新しく作成したりするテクニックを説明します。

6.1. 既存のワークフローを開く

既存のワークフローを開くには:

1. メニューバーから「File(ファイル)」→「Open File...(ファイルを開く...)」を選びます。標準ファイルダイアログボックスが表示されます。
2. もしファイルダイアログボックスから Kepler ディレクトリ (Kepler プログラムをインストールした場所) を開けないときは、直接そのディレクトリまで移動します。サンプルワークフローはこのガイドで説明したように Kepler をインストールしたディレクトリの "/demos/getting-started" ディレクトリ*訳注に保存されています(図7)。



3. ワークフローファイルはダブルクリックによって開くことができます。ワークフローはアプリケーションウィンドウのワークフローキャンバスへ表示されます。

*訳注 "/"はディレクトリ(フォルダ)の区切り文字。日本語版 Windows では"¥"。Kepler(ver.1.0.0)をインストールしたのであれば、デフォルトで"C:¥Program Files¥Kepler-1.0.0"

6.1.1. 例 1：Lotka-Volterra(ロトカ・ヴォルテラ) ワークフロー

この例ではある特定のワークフロー、古典的な捕食者-被捕食者モデルであるロトカ・ヴォルテラ ワークフローを開くことにします。このワークフローを開くには:

1. メニューバーから、「File(ファイル)」→「Open File...(ファイルを開く...)」を選ぶ。標準ファイルダイアログボックスが表示されます。
2. Kepler をインストールしたディレクトリの下の"/demos/getting-started"ディレクトリへ移動すると"02-LotkaVolterraPredatorPrey.xml"があります(図7)
3. "02-LotkaVolterraPredatorPrey.xml"をダブルクリックします。ロトカ・ヴォルテラワークフローは新しく開いたアプリケーションウィンドウのワークフローキャンバスへ表示されます(図8)。

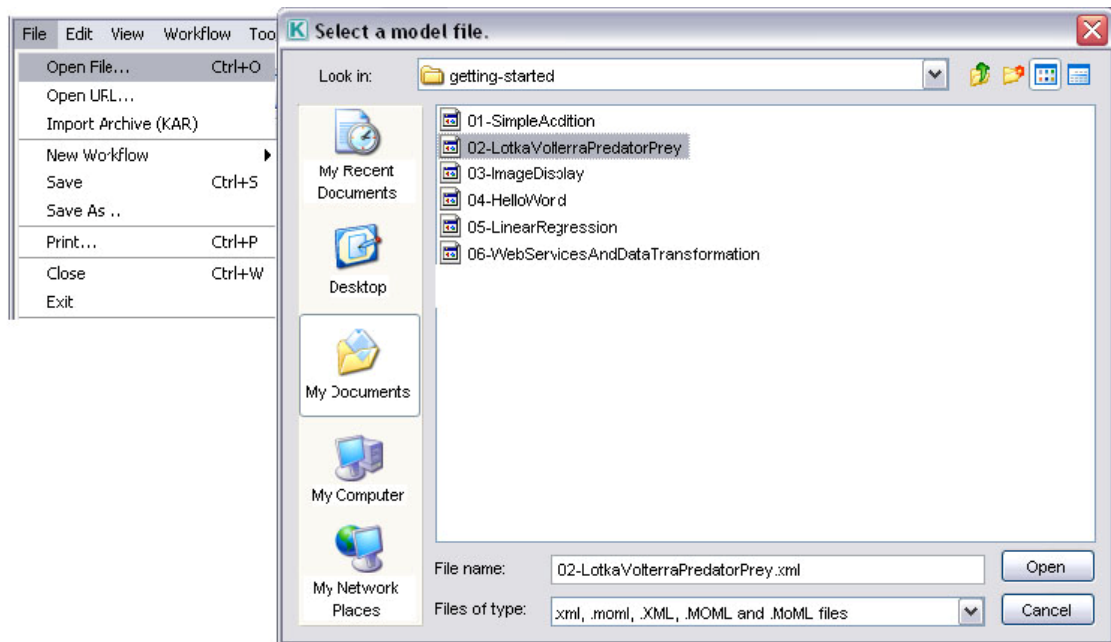


図 7: ロトカ・ヴォルテラ ワークフローへ移動。このワークフローは Kepler がインストールされたディレクトリの中の"/demos/getting-started"ディレクトリにある

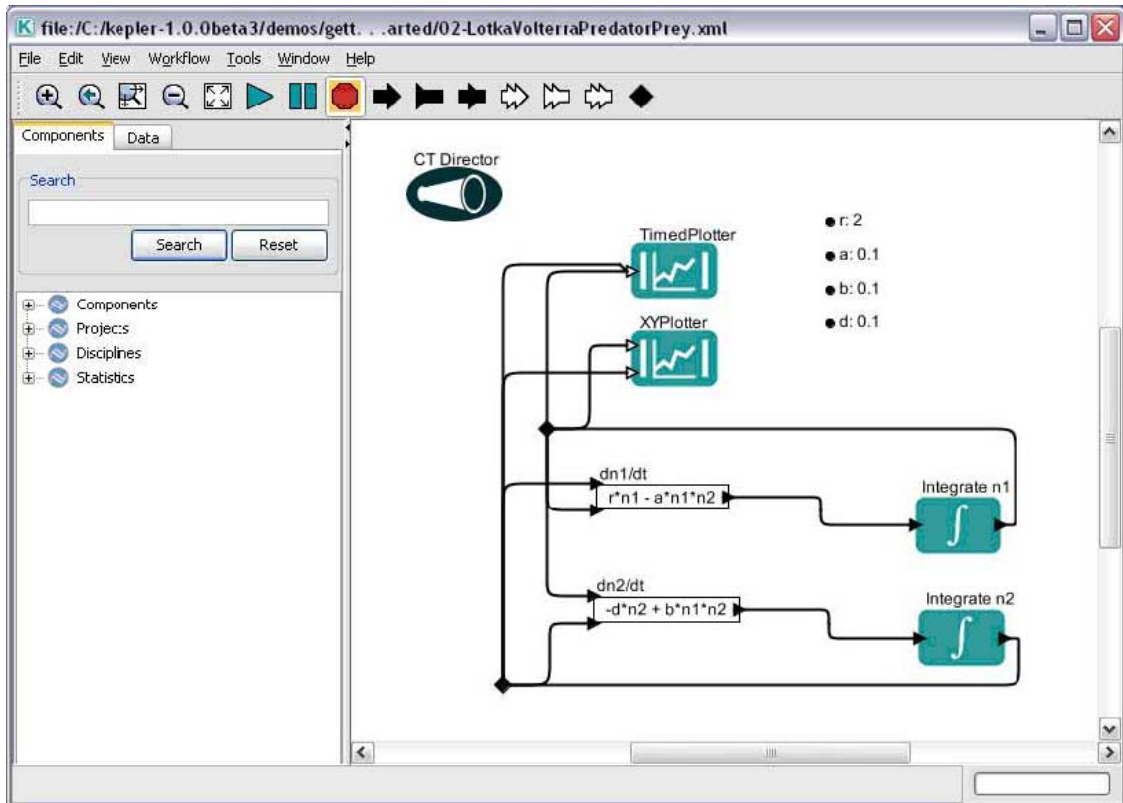



図 8: Kepler インターフェース内のロトカ・ヴォルテラ ワークフロー

6. 2. 既存の科学ワークフローの実行

既存の科学ワークフローを実行するには:

1. 必要なワークフローを開きます。
2. ツールバーから「Run(実行)」  ボタンを選びます。
3. ワークフローが実行され、その出力結果を生成します。

あるいは

1. 必要なワークフローを開きます。
2. メニューバーから「Workflow(ワークフロー)」を選び、その後「Runtime Windows(実行ウィンドウ)」を選びます。実行ウィンドウ(図9)が表示されます。ワークフローがパラメーターを持つ場合は、このウィンドウに表示されます。
3. 必要であればパラメーターを修正し、その後「Go(実行)」ボタンを押します。
4. ワークフローが実行され、その結果を生成します。ワークフロー実行中は、「Pause(一時停止)」、「Resume(再開)」、「Stop(停止)」ボタンを押すことができます。

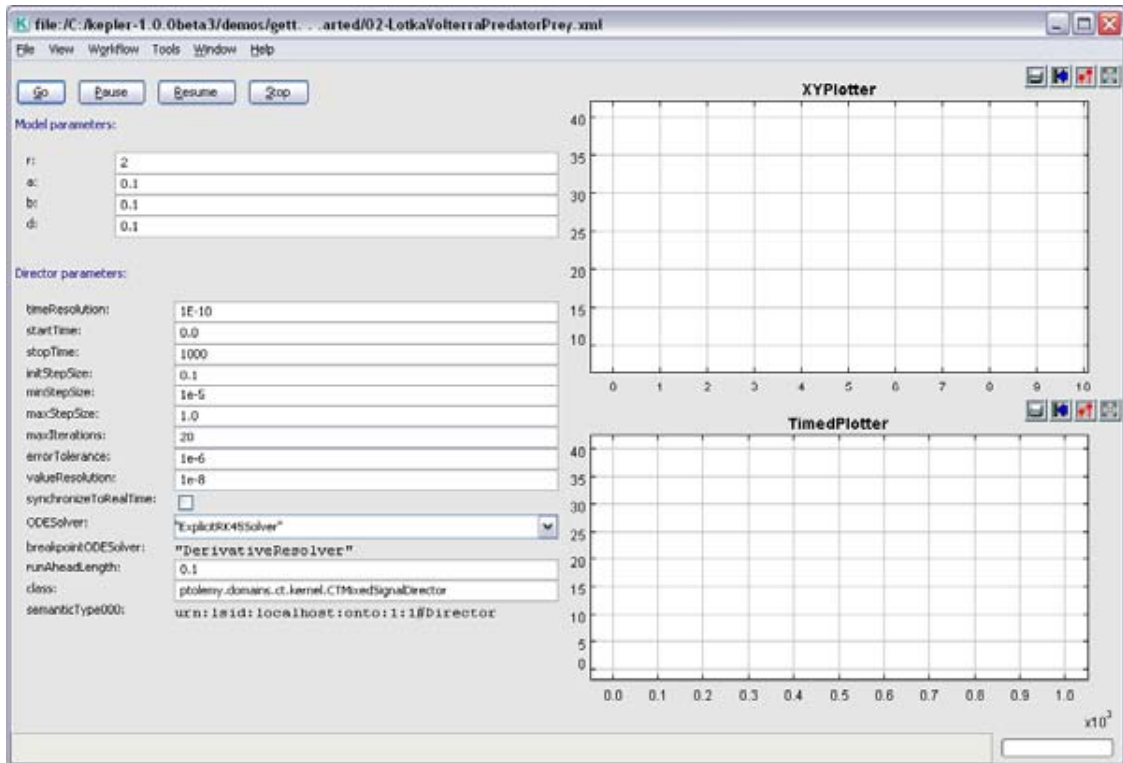


図 9: ロトカ・ヴォルテラ ワークフローを表示させた際の実行ウィンドウ。「Go(実行)」ボタンを押すとワークフローが実行される。ディレクターとモデルパラメーターは実行ウィンドウで変更できる。出力も同様にこのウィンドウへ表示される

6.2.1. 例 2: デフォルトパラメーターでロトカ・ヴォルテラ ワークフローを実行する

ロトカ・ヴォルテラモデルは、Kepler の中で 2 つの連結した微分方程式を解くため、連続時間領域(つまり CT ディレクター)を使います。方程式の一つは捕食者の個体数、もう一つは被捕食者の個体数をモデルにしています。これらを計算すると結果がプロットされ、個体数の変化と位相図の両方が示されます。このモデルの詳細については、セクション 6.2.2 を参照して下さい。

ロトカ・ヴォルテラ ワークフローを実行するには:

1. Kepler をインストールしたディレクトリの下での"/demos/getting-started"ディレクトリの、"02-LotkaVolterraPredatorPrey.xml"を開きます。
2. ツールバーから「Run(実行)」ボタンを選びます。
3. ロトカ・ヴォルテラ ワークフローをデフォルトパラメーターで実行すると 2 つのグラフが作成されます。TimedPlotter というグラフは、時間経過にしたがった捕食者数と被捕食者数との相互の関係を表示します(つまり、モデルで予測される捕食者数と被捕食者数の周期的な変化)。XYPlotter というグラフは個体群の周期的な変化の位相図を表示します(つまり被捕食者数が縦軸で、捕食者数が横軸)。

この2つのグラフはともに捕食者数と被捕食者数とがどのように関連しているかを示しています: 被捕食者が増えると、捕食者も増加します(図10)。

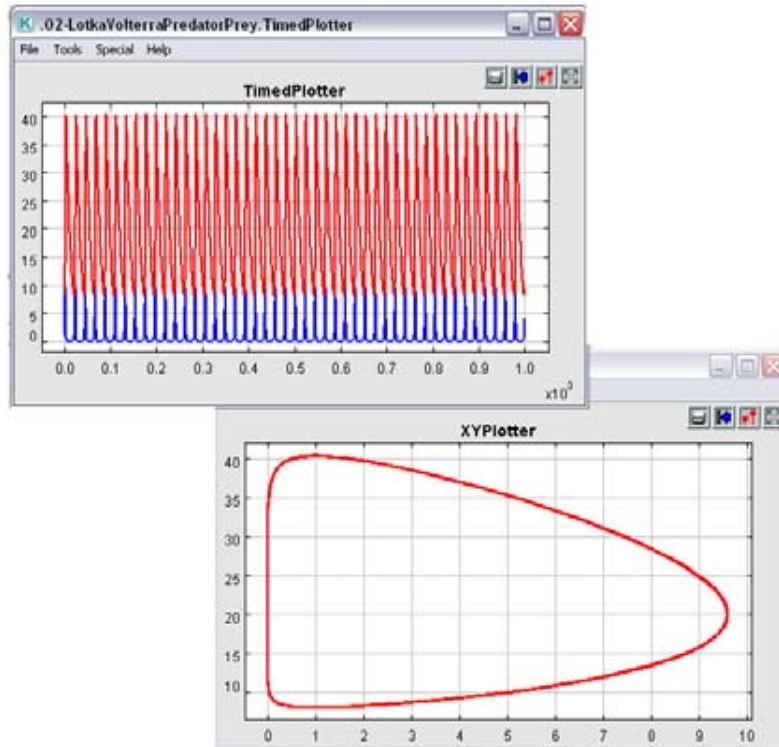


図10: ロトカ・ヴォルテラ ワークフローによるグラフ出力

6.2.2. 例2: 修正したパラメーターでロトカ・ヴォルテラ ワークフローを実行する

ワークフローに対するパラメーターの効果を示すため、ロトカ・ヴォルテラ ワークフローについての背景を少し説明する必要があります(図11)。

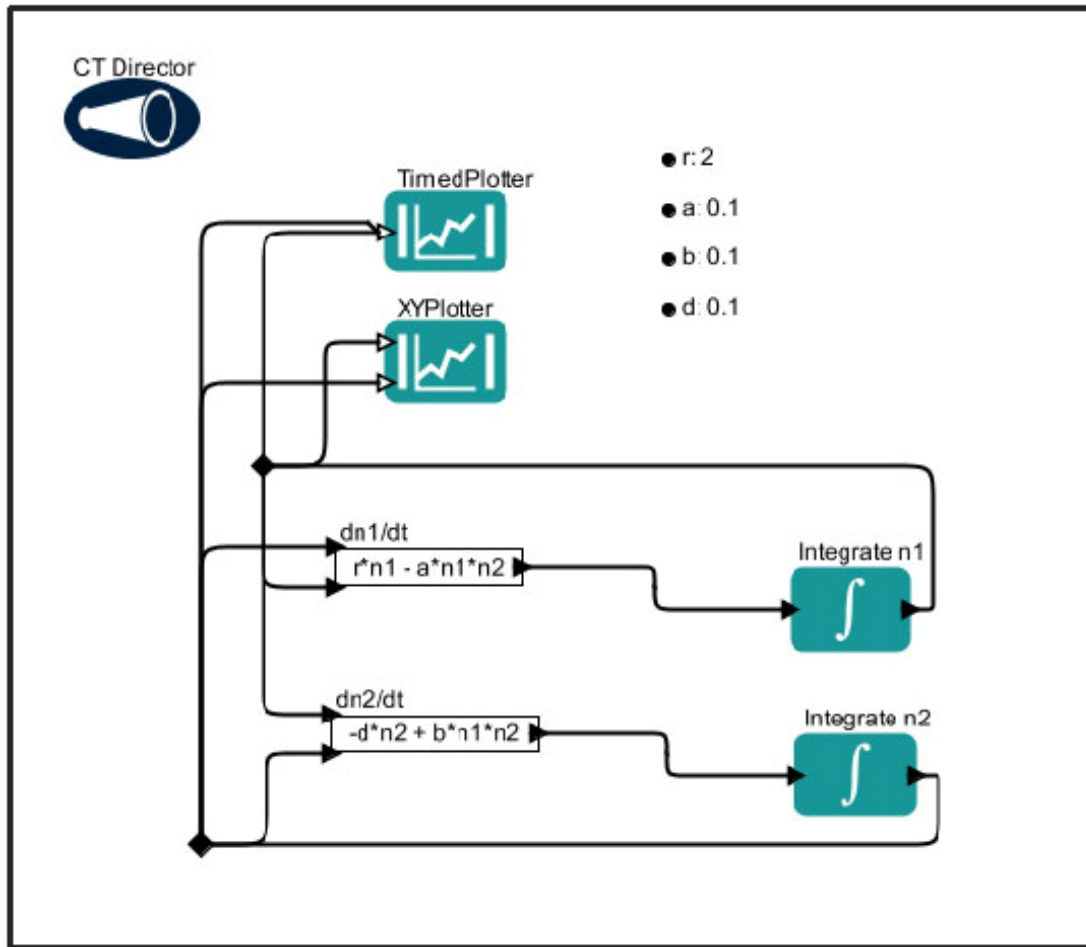


図 11: ロトカ・ヴォルテラ ワークフロー

ロトカ・ヴォルテラ モデルは Lotka(1925)²と Volterra(1926)³によって個々に開発され、2つの微分方程式からなります。片方は被捕食者の個体数の変化を記述し($dn1/dt = r*n1 - a*n1*n2$)、第2の式は捕食者の個体数の変化を記述しています($dn2/dt = -d*n2 + b*n1*n2$)。

ロトカ・ヴォルテラモデルは次のような仮定にもとづいています。

- 被捕食者は無限の資源(餌や生息地など)を持っている
- 被捕食者の唯一の脅威は捕食者である
- 捕食者は餌についてスペシャリストである(つまり捕食者の唯一の餌は被捕食者である)
- 捕食者の成長は捕獲した被捕食者に依存している

Kepler の中で科学ワークフローとして表現されるロトカ・ヴォルテラ モデルは次を含みます:

² Lotka, Alfred J (1925). Elements of physical biology. Baltimore: Williams & Williams Co.

³ Volterra, Vito (1926) Fluctuations in the abundance of a species considered mathematically. Nature 118. 558-560.

- 6つのアクター: 2つの Plotter(プロッター)、2つの Equation(式)、2つの Integral function(積分関数)
- 1つのディレクター
- 4つのワークフローパラメーター(表3)

注意: ロトカ・ヴォルテラ モデルのディレクターも、2つのプロッターアクターと同じように設定可能なパラメーターを持ちます

上述した重要な仮定がワークフローパラメーターの基礎となります。ワークフローパラメーターとそのデフォルト値は次のようになります:

パラメーター	デフォルト値	説明
r	2	捕食者がいない場合の被捕食者の固有成長率
a	0.1	捕食者の捕獲効率、あるいは補食による被捕食者の死亡率
b	0.1	消費された被捕食者バイオマスが捕食者バイオマスに転換される割合(つまり補食者が新しい被捕食者へと変わる効率)
d	0.1	捕食者の死亡率

表 3: ロトカ・ヴォルテラ ワークフローのデフォルトパラメーター

このワークフローで使われている微分方程式($dn1/dt = r*n1 - a*n1*n2$) と ($dn2/dt = -d*n2 + b*n1*n2$)のうち、変数 $n1$ は被捕食者密度、 $n2$ は捕食者 密度を表します。ワークフローのパラメーターを変更する際には、モデルの仮定に注意する必要があります。例えば、もしウサギが被捕食者でキツネが捕食者であるロトカ・ヴォルテラ モデルを作りたいなら、キツネの個体群の動きに対応して、どのようにウサギの個体群が変わるかということについて、次のような仮定をすることができるでしょう:

- ウサギの個体群は被捕食者の制限がなければ指数関数的に成長する
- ウサギの死亡率はキツネの捕獲によって決まる
- キツネは、出会った数に比例してウサギを食べる
- キツネの個体群は食べたウサギの数と食べられたウサギが新しいキツネの子供に変わる効率によって決まり、
- キツネの死亡率は自然の過程によって決まる

もしこれらのプロセスが生ずる速度に着目してモデルを実行したいなら、一定期間における変化の割合を決めるパラメーターを変えてみようと思いませんか。

修正したパラメーターでロトカ・ヴォルテラ ワークフローを実行するには

1. Kepler をインストールしたディレクトリの下での"/demos/getting-started"ディレクトリの、"02-LotkaVolterraPredatorPrey.xml" という名前のワークフローを開きます
2. メニューバーから「Workflow (ワークフロー)」を選び、その後「Runtime Window (実行ウィンドウ)」を選びます。実行ウィンドウが表示されます。2つのパラメーターセットがあることに注意してください一つはワークフローで

(Model parameters)もう一つはディレクターです(Director parameters)。この例では、両方のパラメーターセットを修正することにします。

3. 表 4 に示したように、ワークフローパラメーターを変更して下さい。

パラメーター	値	説明
r	0.04	捕食者がいない場合の被捕食者の固有成長率
a	0.0005	捕食者の捕獲効率、あるいは補食による被捕食者の死亡率
b	0.1	消費された被捕食者バイオマスが捕食者バイオマスに転換される割合 (つまり捕食者が新しい被捕食者へと変わる効率)
d	0.2	捕食者の死亡率

表 4: ロトカ・ヴォルテラ ワークフローのパラメーター変更。

<http://www.stolaf.edu/people/mckelvey/envision.dir/lotka-volt.html> より

4. ディレクターパラメーター `stopTime` を 300 に修正して下さい。

5. 実行ウィンドウにある「Go(実行)」ボタンをクリックして下さい。

ロトカ・ヴォルテラ ワークフローを修正したパラメーターで実行すると 2 つのグラフが作成されます。TimedPlotter グラフと XYPlotter グラフです。パラメーターの変化によって、捕食者と被捕食者の数はやはり連動していますが、その関係性は変化しています(図 12、デフォルトパラメーターの図 10 と比較して)。

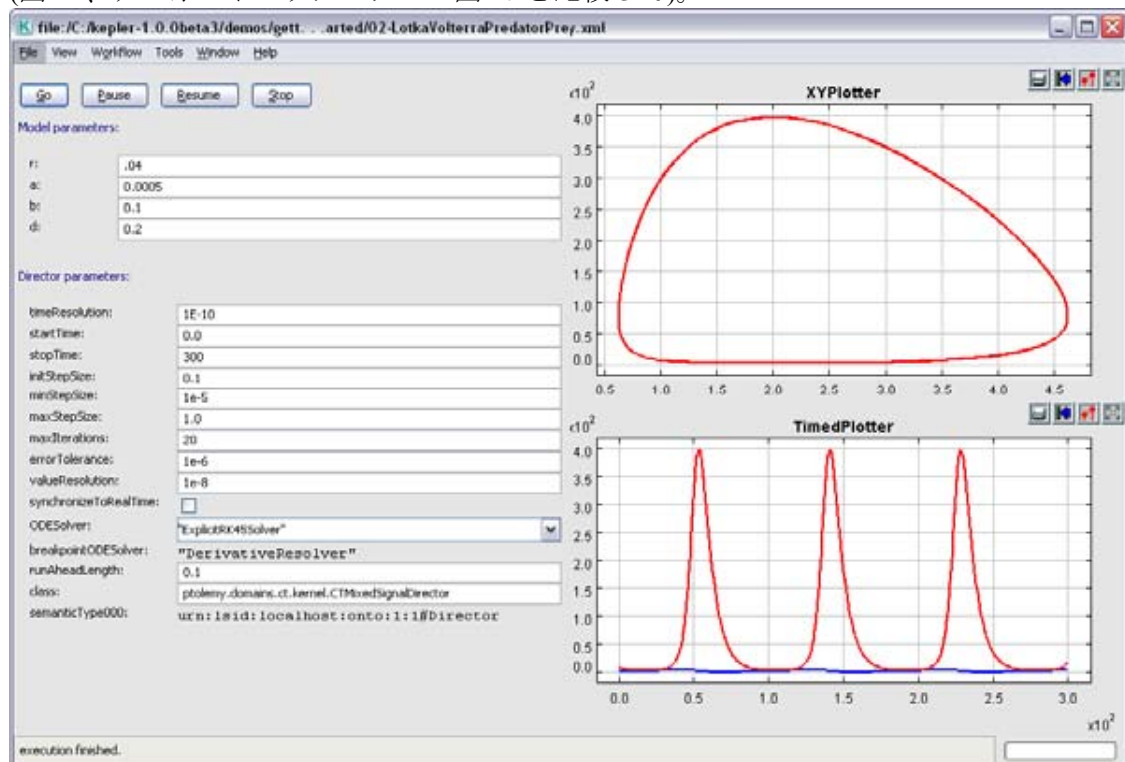


図 12: 修正パラメーターでのロトカ・ヴォルテラ モデルによるグラフ出力

6.3. 既存の科学ワークフローを編集する

既存の科学ワークフローを編集する方法には二つあります。

1. 現在のデータセットを別のデータセットに置き換える、もしくは
2. ワークフローの中のいくつかの解析プロセスを別な解析プロセスに置き換える
(例：ニューラル・ネットワーク・モデル・アクターを確率モデルアクターに置き換えるなど)

データやプロセスを置き換える前に、それらのアクターが要求するインプットとアウトプットが何かを理解してはなりません。

注意：アクターの細かい説明を見るには、そのアクターの上で右クリックし、「Documentation」、次に「display」を選択してください。(図13)。そのアクターの主な機能と要求されるインプットとアウトプットを説明するダイアログボックスが開きます。ダイアログを見終わったらそのウィンドウを閉じてください。

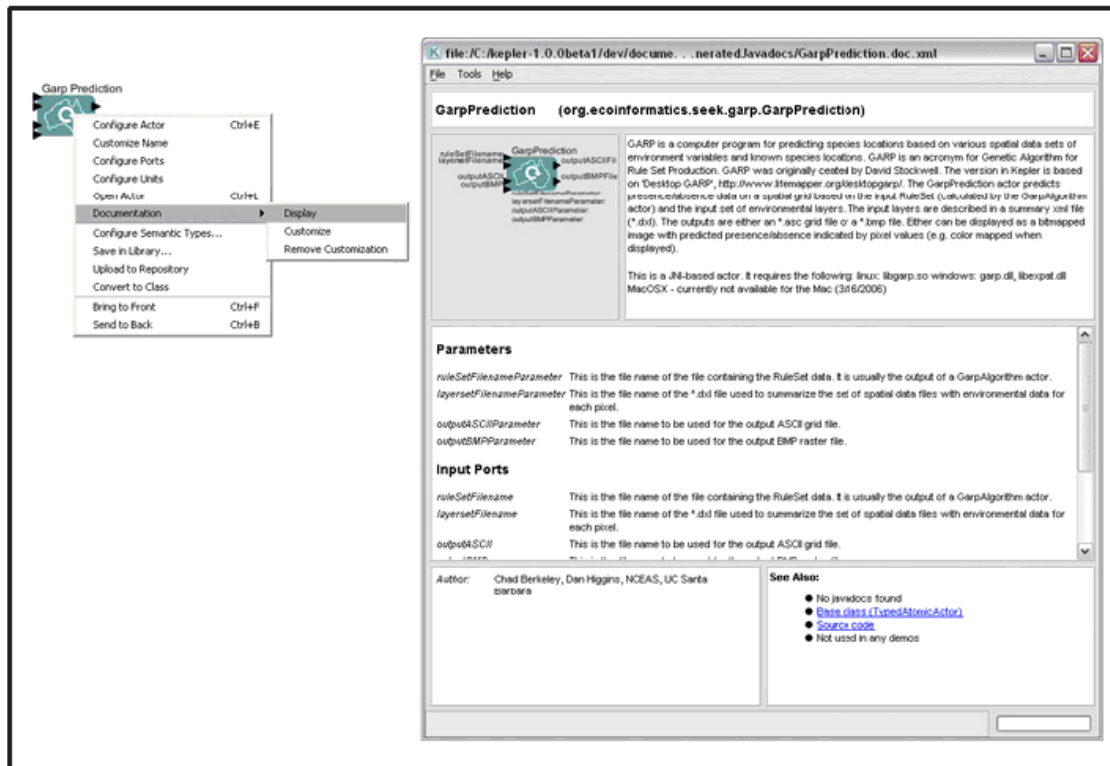


図 13: アクターの解説の表示

既存の科学ワークフローを編集するには、

1. 使いたいワークフローを開く。
2. 置き換えるワークフローコンポーネントを決める。
3. 対象のコンポーネント（データアクターかプロセスアクター）をクリックして選択する。選択されたコンポーネントは濃い黄色の境界線で強調される。

4. キーボードの削除キーを押す。強調されたコンポーネントがワークフローキャンバスから消える。
5. コンポーネント・データアクセスエリアから、適切なデータアクターかプロセスアクターをワークフローキャンバスにドラッグする。
6. 適切なインプットとアウトプットポートを接続する。
7. ワークフローを実行する。
8. メニューバーから「File(ファイル)」、つぎに「Save(保存)」（既存のワークフローを上書きする場合）を選ぶか、「Save As...(名前を付けて保存)」（新しいワークフローとして保存する場合）を選ぶ。「名前を付けて保存」を選択した場合は、指示に従って新しい名前を入力する。

6.3.1. 例 4: Image J ワークフローの解析プロセスを編集／置き換える

この例では、二つの異なるアクターがどのように同じ機能を果たすのかを示します。Kepler の “/demos/getting-started” フォルダにあるイメージ・ディスプレイ・ワークフロー (Image Display workflow)(03-ImageDisplay.xml) を使い、*ImageJ* アクターを *BrowserDisplay* アクターに置き換えます。これらのアクターは二つとも、Mephitis 属（スカンクの一属名）の南北アメリカにおける分布を表すビットマップ画像を表示します（この画像は GARP という遺伝的アルゴリズムにより作成されました。GARP はある種が個体群を維持することのできる環境条件を示す生態学ニッチモデルを作ることができます。GARP は [サンディエゴ・スーパー・コンピューター・センター](http://www.lifemapper.org/desktopgarp/) の David Stockwell により開発されました。GARP についてもっと知りたい方は、<http://www.lifemapper.org/desktopgarp/> をご覧ください）。

イメージ・ディスプレイ・ワークフローを編集するには、

1. Kepler の “/demos/getting-started” ディレクトリから 03-Image-Display.xml ワークフローを開きます。
2. ここでの置き換え対象となる *ImageJ* アクターを選択します。*ImageJ* アクターが選択されていることを示す濃い黄色の境界線で強調されます。（図 14）

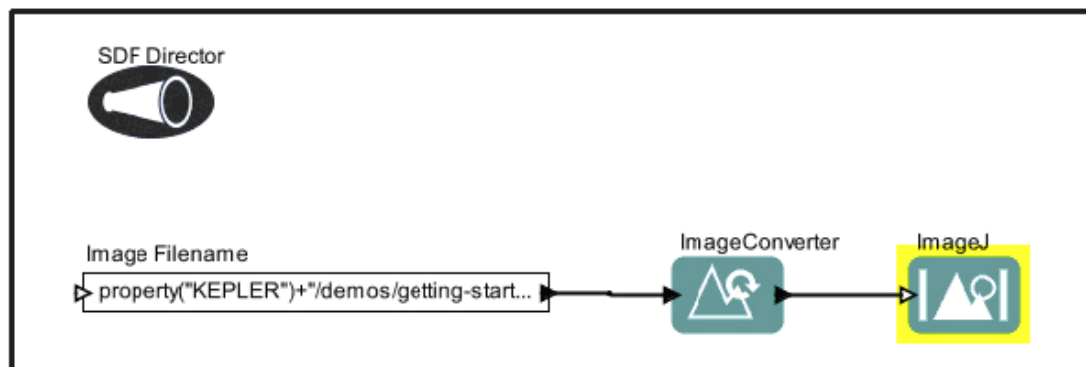


図 14: *ImageJ* アクターが強調表示されている Image Display ワークフロー

3. キーボードの削除キーを押します。 *ImageJ* アクターがワークフローキャンバスから消えます。
4. データアクセスエリアから *Browser Display* アクターをワークフローキャンバスへドラッグします。 *Browser Display* アクターは、「コンポーネント」タブの「Components > Data Output > Workflow Output > Textual Output」にあります。
5. *Image Converter* アクターのアウトポートポートを *Browser Display* アクターのインポートポートに接続します。ポートをつなぐには、 *Image Converter* アクターの右側にあるアウトポートポート（黒い三角）を左クリックして、マウスボタンを押し下げたまま、ポインタを *Browser Display* アクターの左側上部のインポートポートにドラッグし、マウスを放します。接続が完了したら、太い黒い線が表示されます。もし接続がうまく行かなかったら薄い線が表示されます。
6. ワークフローを実行します。
7. メニューバーから「File (ファイル)」、つぎに「Save (保存)」（既存のワークフローを上書きする場合）を選ぶか、「Save As (名前を付けて保存)」（新しいワークフローとして保存する場合）します。「Save As」を選択した場合は、指示に従って新しい名前を入力します。

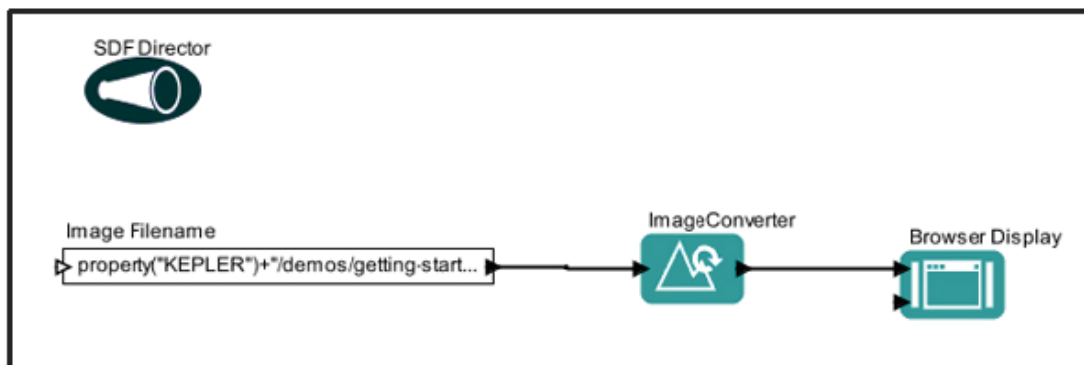


図 15: *ImageJ* アクターが *Browser Display* に置き換えられた *Image Display* ワークフロー。

注意: 多くの場合、アクター同士を接続するには接続元のアウトポートポートから接続先のインポートポートに向けてつなぐのが一番簡単です。

6. 4. Kepler で検索する

Kepler にはデータ (EarthGrid 上にあるもの) や解析プロセスコンポーネント (ローカルシステムのみ、もしくはローカルシステムとリモートコンポーネントリポジトリの両方にあるもの) を見つけるための検索メカニズムがあります。このセクションの例では Kepler でデータとコンポーネントを検索する方法について説明します。

6. 4. 1. 利用可能なデータを検索する

検索機能により、Kepler は EarthGrid のデータへのアクセスを可能にします。EarthGrid のリソースは KNB Metacat <http://knb.ecoinformatics.org> か、カンザス大学の Digir <http://www.specifysoftware.org/Informatics/informaticsdigir/> に置かれています。Kepler から EarthGrid を検索するには、

1. コンポーネント・データアクセスエリアから「データ」タブを選択 (図16) します。
2. 好きな検索キーワードをタイプします (例: **Datos Meteorologicos** など)。検索キーワードのスペルが正しいことを確認してください (検索キーワードは言葉の一部でも入力可能です - e.g. 'Datos' など)。
3. **Search** (検索) ボタンをクリック。検索には数秒かかる場合があります。ログインを求められる場合もあります。もし求められたら、あなたのユーザとパスワードを入力するか、"**Login Anonymously** (匿名でログイン)" をクリックしてください。検索が完了すると、コンポーネント・データアクセスエリアに検索結果 (つまり、データアクター) のリストが表示されます。
4. 検索結果の中のデータアクターを使うには、使いたいアクターをワークフローキャンバスにドラッグします。

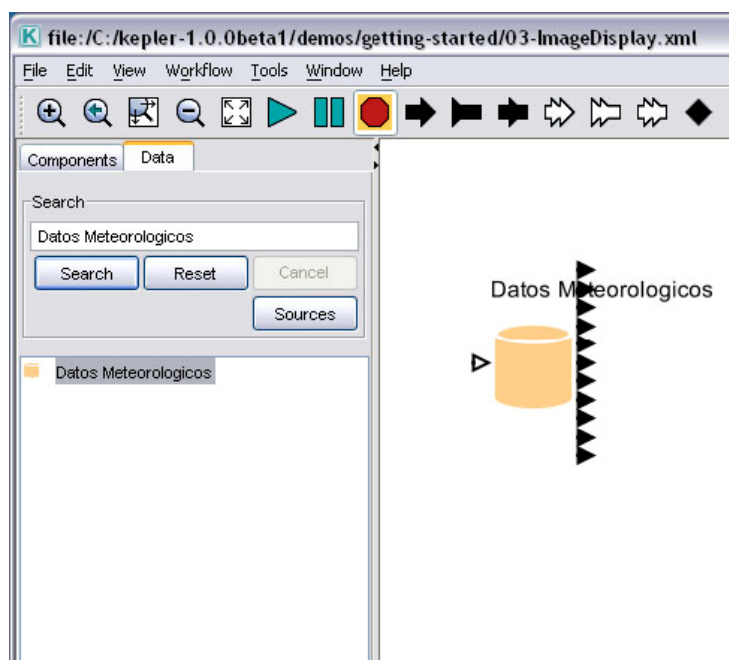


図 16: Datos Meteorologicos を見つけるための検索

注意: データ検索の設定を変えるには、**Sources** (ソース) ボタンをクリックします。検索対象のソースと、読み出すドキュメントのタイプを選択します。

データアクターの情報は三つの方法で表示することができます。(1) ワークフローキャンバスで、データアクターのアウトポートポートの上にマウスを移動するとそれぞれのポートの名前とデータアウトプットのタイプを示すツールチップを表示します；(2) データアクターを右クリックし、「**Get Metadata**」を選択すると、そのデータセットに関するさらなる情報を含むウィンドウが開きます；(3) データアクターを右クリックし、ドロップダウンメニューから「**Preview**」を選ぶとデータセットをプレビューすることができます (図17)。

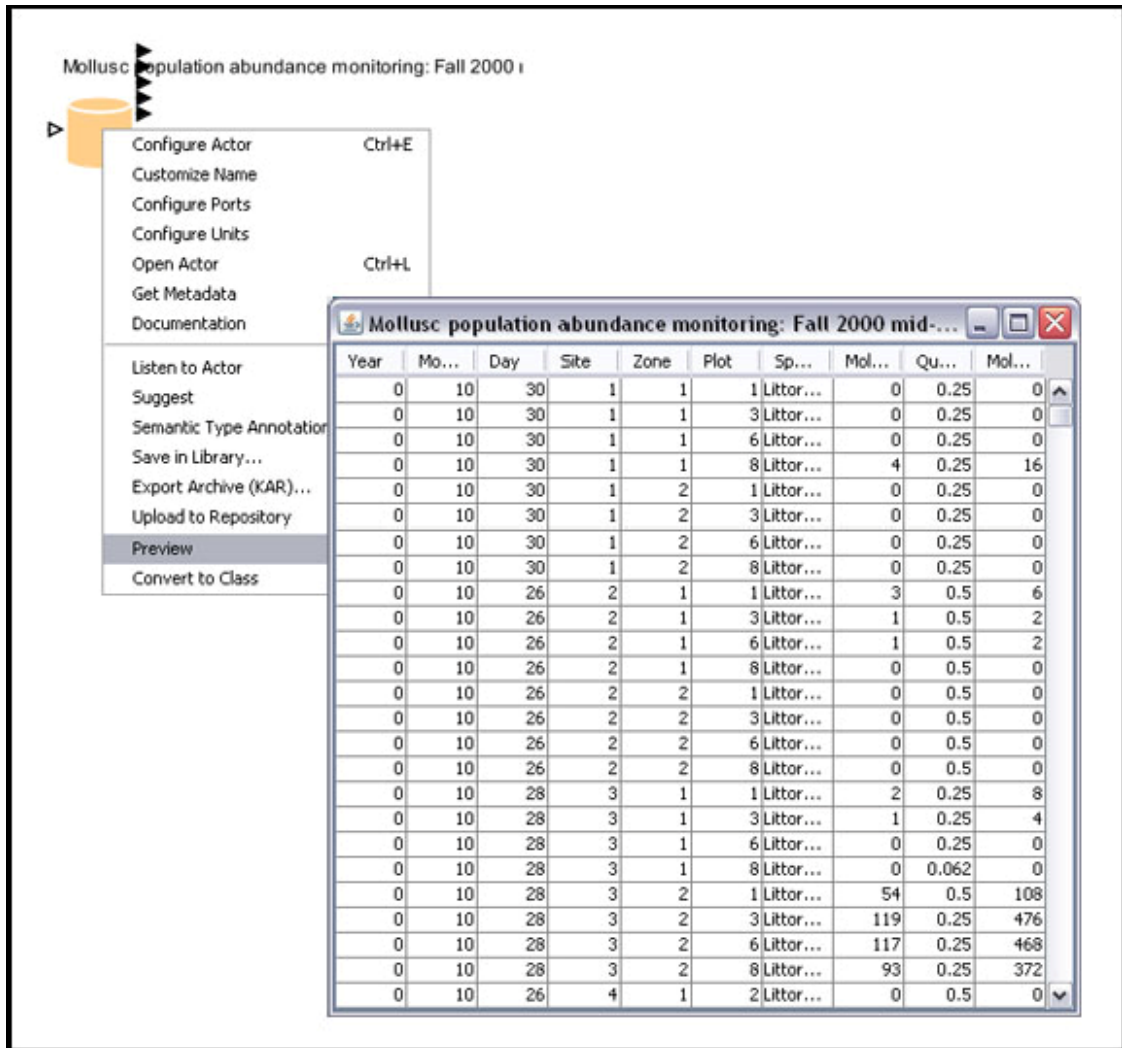


図 17: データセットのプレビュー

6. 4. 2. 利用可能なプロセッシング・コンポーネントを検索する

Kepler には 350 以上のワークフローコンポーネントが付属しています。様々な解析機能を用いていくつでもワークフローを作ることができます。デフォルトの Kepler プロセッシングコンポーネントはコンポーネント・データアクセスエリアの「コンポーネント」タブの下に表示されています。コンポーネントは機能によって整理されています（例：“Director（ディレクター）”、“Filter Actor（フィルターアクター）”など）。コンポーネントを検索するには、

1. ワークフローキャンバスの左側にあるコンポーネント・データアクセスエリアで「コンポーネント」タブを選択します。
2. 好きな検索キーワードをタイプします（例：“File Fetcher”など）。
3. 「Search」ボタンをクリック。検索が完了すると、コンポーネント・データアクセスエリアに検索結果のリストが表示されます。検索結果はデフォルトのコンポーネントリストに書き換えられます。（なぜなら、コンポーネントはカテゴリー

毎に並べられているので、同じコンポーネントが検索結果の複数の場所に現れるかもしれないからです)。

4. 検索結果の中のコンポーネントを使うには、使いたいコンポーネントをワークフローキャンバスにドラッグします。
5. 検索結果を削除し、デフォルトのコンポーネントを表示させるには、「Reset」ボタンをクリックします。

注意：もし使いたいコンポーネントとそのコンポーネントの場所が分かっていたら、直接ライブラリから探してそのコンポーネントをワークフローキャンバスへドラッグできます。

6.5. 基本的な科学ワークフローを作成する

あなた自身の実行可能ワークフローをデザインし、作成し、保存できるのが Kepler の特長の一つです。ワークフローを作成する一般的な順序は以下の通りです。

1. あなたの科学ワークフローの概念モデルを（紙やその他の媒体に）作ります。
2. Kepler を起動します。
3. Kepler で利用可能なデータとアクターコンポーネントをあなたの概念モデルに配置します。
4. あなたのワークフローに適したディレクターを選び、ワークフローキャンバスにドラッグします。ディレクターの選び方についてのより詳しい説明は Kepler ユーザーマニュアルの 5 章をご覧ください。
5. 使いたいワークフローコンポーネントをワークローキャンバスにドラッグします。
6. ワークフローコンポーネントを接続します。
7. ワークフローを保存します。

このセクションの例ではあなた自身のワークフローを作るための初歩を解説します。最初の例では、古典的な“Hello World”(ハローワールド)ワークフローを使って、Kepler で実行可能なワークフローを作るのがいかに簡単かをお見せします。二つ目の例ではより実践的に、あなたのデスクトップ上のデータをワークフローに取り込む方法をお見せします。

6.5.1. 例 5: “Hello World (ハローワールド)” ワークフローを作成する

“ハローワールド”ワークフローを作るために、まず使用するデータのタイプ（例：テキスト、文字列データなど）、望ましいアウトプットのタイプ（テキスト形式、イメージ表示など）、または、このモデルを実行するのに必要なディレクター（同期、または並列）、などについて考えるところから始めましょう。“ハローワールド”ワークフローには *Constant* (定数)アクター、*Display*(テキスト出力)アクター、そして *SDF* ディレクター（初期値では、*SDF* ディレクターの中では、データはワークフロー中のアクターの順番に従って処理され、そのワークフローは連続的に実行されます）が必要です。

1. Kepler を起動します。空白のキャンバスが開きます。
2. コンポーネント・データアクセスエリアで「コンポーネント」タブを選択し、「/Components/Director/」ディレクトリを開きます。
3. SDFディレクターをワークフローキャンバスの上部にドラッグします。
4. さらに「コンポーネント」タブで“Constant”を探し、Constantアクターを選択します。
5. Constantアクターをワークフローキャンバスにドラッグし、SDFディレクターの少し下に配置します。
6. Constantアクターを右クリックし、メニューから「Configure Actor」を選んで設定します。(図18)

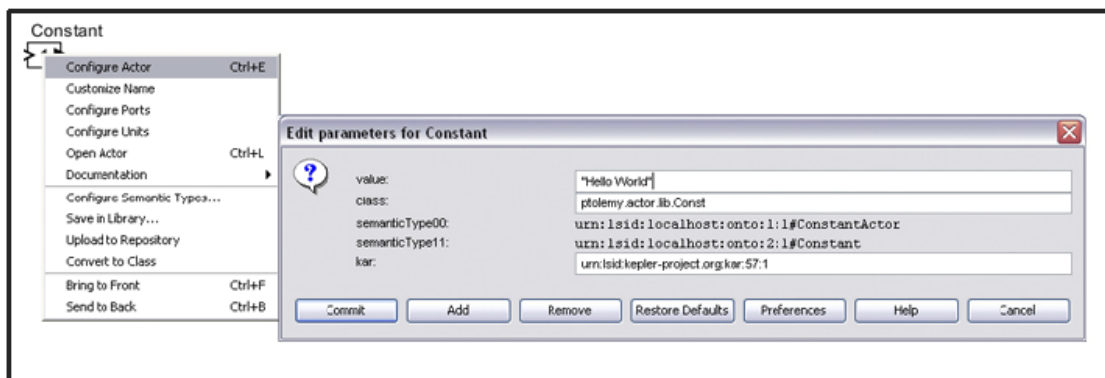


図 18: Constant アクターの設定

7. “Edit parameters for Constant”ダイアログウィンドウの value (値) フィールドに “Hello World”とタイプし、変更を保存するために「Commit (承認)」をクリックします。“Hello World”は文字列値です。Kepler では、すべての文字列値は引用符で囲わなくてはなりません。
8. コンポーネント・データアクセスエリアで“Display”を検索し、“Textual Output”の下での Displayアクターを選択します。
9. Displayアクターをワークフローキャンバスにドラッグします。
10. Constantアクターのアウトポートポートを Displayアクターのインポートポートに接続します。
11. モデルを実行します。(図19).

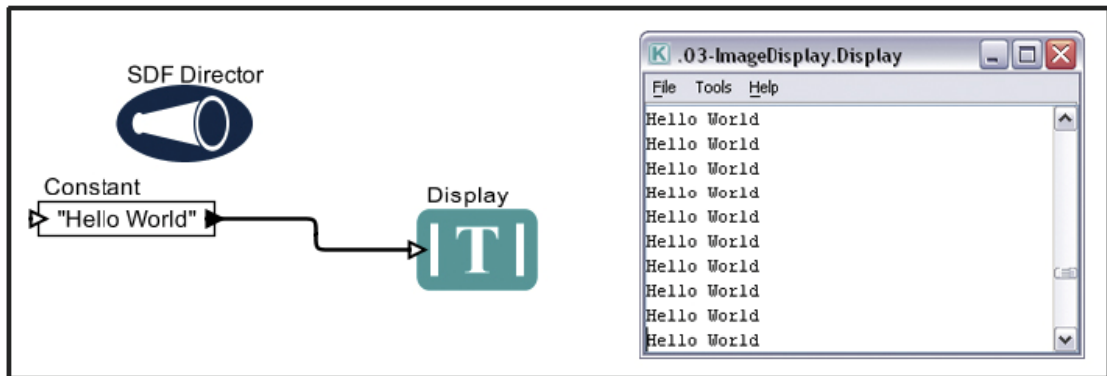


図 19: “Hello World” ワークフローとアウトプット

注意: 初期値では、*SDF* ディレクターはワークフローのループを作って連続的に実行します。“Hello World” ワークフローを決められた回数だけ実行したい場合は、*SDF* ディレクターを右クリックして、メニューから“Configure Director”を選択してください。“Edit parameters for SDF Director”ダイアログの *iterations* フィールドに希望の回数をタイプし、変更を保存するために *Commit* をクリックします。

6.5.2. 例 6: 自分のデータを使って簡単なワークフローを作成する

この例では、種数についての情報の入ったローカルのデータファイルを読み込むアクターを使って簡単なワークフローを作ります。

Kepler は様々な形式のデータを様々な方法で読み込むことができます。この例では、データテーブルを閲覧するアクターを使います。どのアクターを使うのが適切かを決めるためには、データが保存されている形式を考えましょう。この例では、データはテキストフォーマットで保存されています。ですからテーブル形式のデータを読み込むことのできる *File Reader* アクターを使いましょう。このワークフローにはテキストを出力するために二つのアクター、*File Reader* アクターと *Display* アクターが必要です。また、この例では *SDF* ディレクターも必要です。

1. メニューバーから「File」を選び、続いて「New Workflow」、そして「Blank」を選びます。新しいウィンドウが開き、空白のワークフローキャンバスが現れます。
2. コンポーネント・データアクセス エリアで「コンポーネント」タブを選択し、“/Components/Director/” ディレクトリを開きます。
3. *SDF* ディレクターをワークフローキャンバスの上部にドラッグします。
4. 「コンポーネント」タブで検索ボックスに“File Reader”とタイプし、「Search」ボタンをクリックします。
5. *File Reader* アクターをワークフローキャンバスにドラッグします。
6. *File Reader* アクターを右クリックし、メニューから「Configure Actor」を選びます。“Edit parameters for File Reader” ダイアログウィンドウが開きます。
7. *fileOrURL* パラメータの右にある「Browse」ボタンをクリックし、*mollusc_abundance.txt* ファイルを指定します。これらのデータは Kepler と一緒にインストールされており、Kepler の “/demos/getting-started/” ディレクトリに入っています。(図 20).

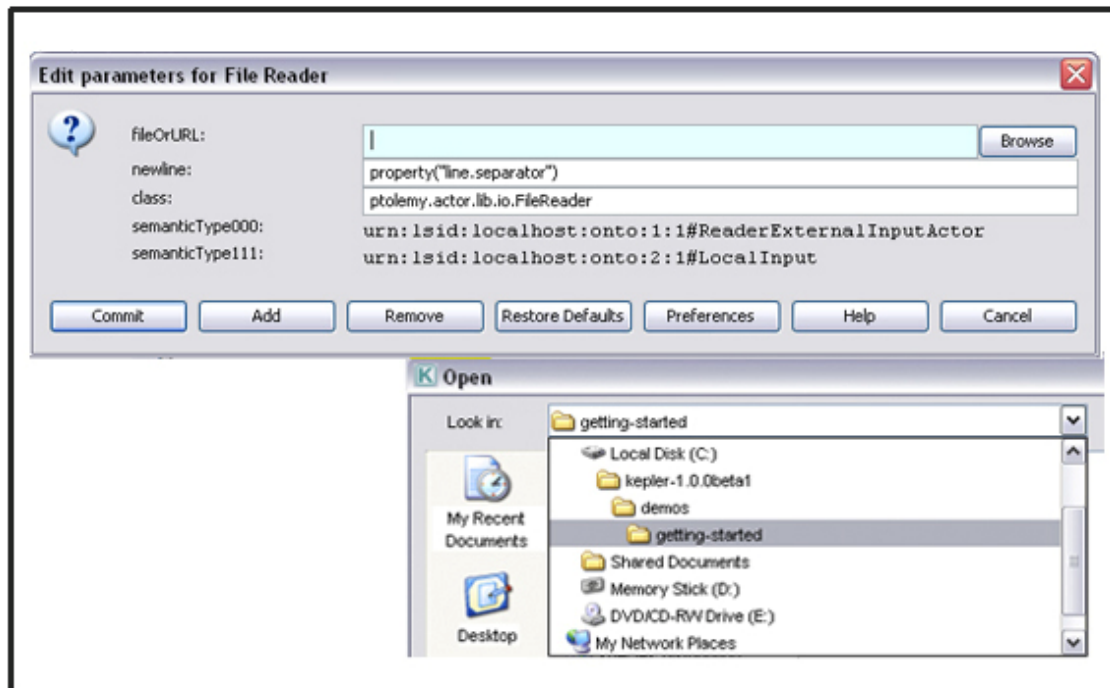


図 20: あなたのローカルコンピュータにあるデータを使うために *File Reader* アクターを設定します

8. “Edit Parameters for File Reader”ダイアログボックスの下部にある「Commit（承認）」ボタンをクリックします。これでこのアクターは特定のファイルを読み込むように設定されました。
9. コンポーネント・データアクセスエリアで“Display”を検索します。*Display* アクターを選択して、ワークフローキャンバスの *File Reader* アクターの右側にドラッグします
10. *File Reader* アクターのアウトポートポートを *Display* アクターのインポートポートに接続します。
11. ツールバーの「Run（実行）」ボタンをクリックします。ポップアップウィンドウが現れ、データファイルの内容をテーブル形式で表示します。(図21)。
12. メニューバーから「File」、そして「Save」を選びます。指示に従って新しいワークフローに“readingdata.xml.”と名前を付け、Keplerの“demos/getting-started”ディレクトリに保存しましょう。

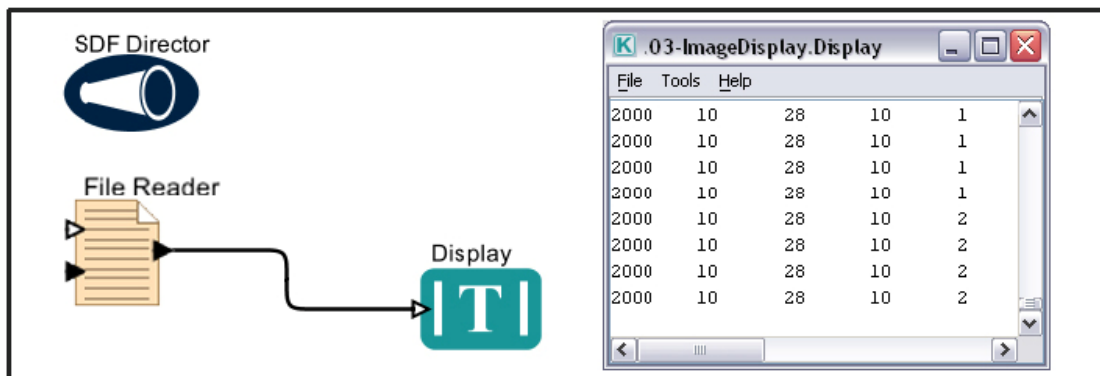


図 21: ローカルデータをワークフローで使用し、表示する

注意：ワークフローを作る時には、データ次第でどのプロセスコンポーネントが使えるかが決まる、と覚えておくとよいでしょう。

7. サンプル科学ワークフロー

このセクションでは、Keplerに標準装備されている小さなサンプル科学ワークフローを詳しく見てみながら、これらのワークフローを作成するための手順を一つ一つご紹介します。

7.1. サンプルワークフロー1 - 基本統計量

名前	基本統計量
ファイル名	/kepler/demos/getting-started/00-StatisticalSummary.xml
詳細	このワークフローはある数値セットの平均、標準偏差、分散を計算します。 <i>Constant</i> アクターにインプットデータが入っています。値の配列{1,2,3,4,5,6,7,8,9,10}です。これらのデータは <i>SummaryStatistics</i> アクターに送られます。 <i>SummaryStatistics</i> アクターは統計の計算が行い、結果をアウトポートから出力します。結果は三つの <i>TextDisplay</i> アクターが表示します。
仮定	この <i>SummaryStatistics</i> アクターは、 <i>RExpression</i> アクターの特 殊版です。 このワークフローを実行するには、 <i>Kepler</i> を実行しているのと同 じコンピュータに、 <i>R</i> (統計計算のための言語・環境) がイン ストールしてあることが必要です。
ディレクター	<i>SDF</i> ディレクター
データ	データは <i>Constant</i> アクターで生成される
アクター	<i>Constant</i> , <i>SummaryStatistics</i> , <i>Display</i>
パラメータ	<i>SDF</i> ディレクター:iterations=1 <i>Constant</i> : value={1,2,3,4,5,6,7,8,9,10}

Summary Statistics(基本統計量)ワークフローは数値のリストを受け取り、平均、分散、標準偏差を計算し、結果を表示します。このワークフローは **Kepler** の機能の使いやすさを示す良い例です。このワークフローを実行するには、**Kepler** を実行しているのと同じコンピュータに、**R** (統計計算のための言語・環境) がインストールしてあることが必要です。**R** は **Windows** 用とマック用の **Kepler** フルインストール版に含まれています。**Linux** 用のインストーラには含まれていません。このワークフローを作成するには、新しいブランクのワークフローを **File** メニュー(**File** > **New Workflow** > **Blank**) から開き、以下の手順に従ってください。

1. コンポーネント・データアクセス エリアで「コンポーネント」タブを選択します。
2. **SDF** ディレクターを検索し、ワークフローキャンバスにディレクターをドラッグします。
3. **SDF** ディレクターを右クリックし、**Configure Director** を選択して設定します。“**Edit Parameters for SDF Director**” ウィンドウで、**iterations** を1に設定して、**Commit** をクリックします。
4. **Constant** アクターを検索し、ワークフローキャンバスにドラッグします。**Constant** アクターは、コンポーネント(**Components**) > ワークフロー(**Workflow**) > ワークフローインプット(**Workflow Input**) > **Constant** から見つけることもできます。
5. **Constant** アクターを右クリックし、**Configure Actor** を選択して設定します。“**Edit Parameters for Constant**” ウィンドウで、**value** フィールドを{1,2,3,4,5,6,7,8,9,10}に設定し、**Commit** をクリックします。注:括弧 {} も必要です。**Kepler** では {} 括弧は配列を指定する時に使われます。
6. **Summary Statistics** アクターを検索し、ワークフローキャンバスにドラッグします。
7. **Summary Statistics** アクターを右クリックして **Configure Ports** を選び、正しい出力ポートを見つけます。(図22)。
8. “**Configure ports for Summary Statistics**” ダイアログボックスで、**Show Name** カラムの **xmean**, **xstd**, **xvar** 行のチェックボックスをクリックします。あなたの行った変更を保存するため、**Commit** をクリックします。これで **xmean**, **xstd**, **xvar** の出力ポートがワークフローキャンバスに現れ、適切なポートに繋ぎやすくなりました。

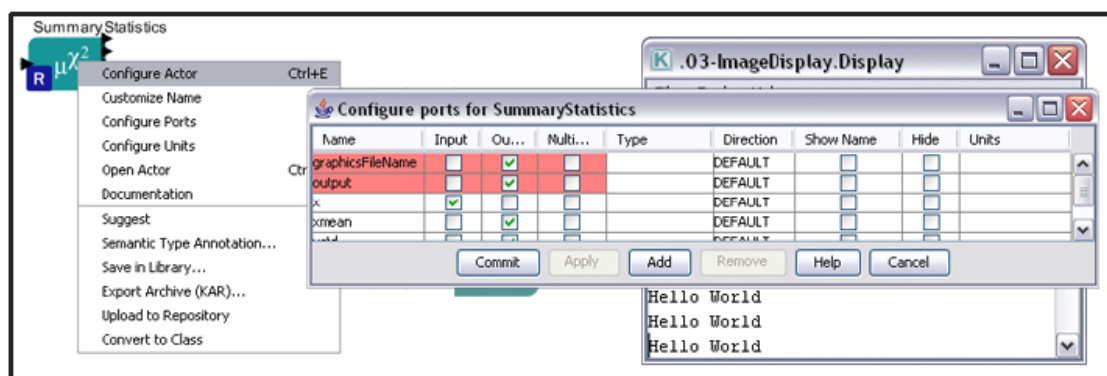


図 22: ポート名を表示する

9. *Constant* アクターの出力ポートを *SummaryStatistics* アクターの出力ポートに接続します。
10. *Display* アクター(テキスト出力)を検索し、ワークフローキャンバスに3回ドラッグします。二つ目のアクターが *Display2*、三つ目のアクターが *Display3* という名前になっていることを確認してください。
11. 三つの *Display* アクターをそれぞれ右クリックし、*Customize Name* を選択して名前を変更します。一つ目の *Display* アクターの“Rename Text Display”ダイアログボックスに“Mean”とタイプし、変更を保存するため、*Commit* をクリックします。同じように、*Display2* の名前を“Variance”に、*Display3* を“Standard Deviation”に変更します。
12. *SummaryStatistics* アクターの *xmean*、*xstd*、*xvar* 出力ポートをそれぞれ対応する *Mean*、*Standard Deviation*、*Variance* アクターに接続します。

これでワークフローを実行する準備ができました。完成したワークフローと出力結果は図23のとおりです。

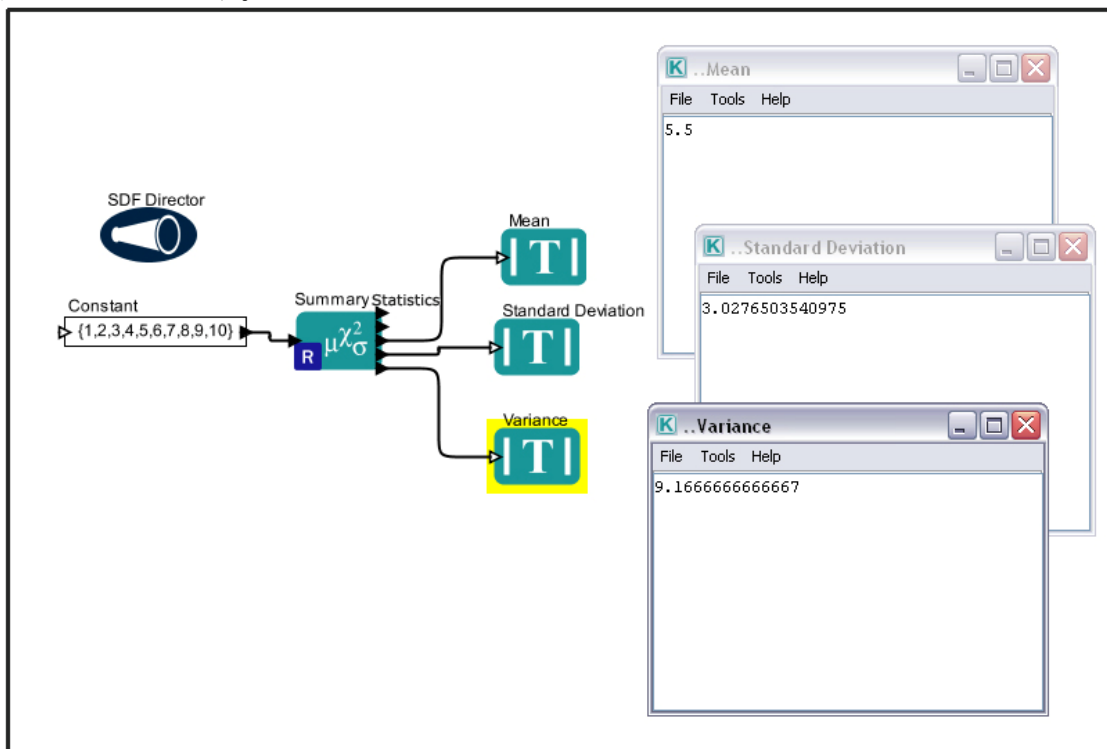


図 23: 基本統計量ワークフローと出力

図 23 の右側のウィンドウには *Constant* アクターの数値配列から作成された平均、分散、標準偏差が表示されています。*Constant* アクターの配列を変えて（例えば、{1,17,6,4,12} などに）、それぞれの統計量を新しく計算してみましょう。

7.2. サンプルワークフロー2 -線形回帰

名前	Rを使った単純線形回帰ワークフロー
ファイル名	/kepler/demos/getting-started/05-LinearRegression.xml
詳細	このワークフローは <i>RExpression</i> アクターを用いて線形回帰解析を実行します。ワークフローは <i>Datos Meteorologicos</i> データセットの二つの変数から散布図を作成し、式 $Y = a + bX$ を使って、 X が説明変数、 Y が従属変数であるような、回帰直線を追加します。 b が回帰直線の傾き、 a が切片 ($X=0$ の時の Y の値) です。
仮定	線形回帰は直線性、独立性、等分散性、正規性を仮定しています。ワークフローを実行するシステムに R がインストールされていなくてはなりません。R はウィンドウズとマック用の Kepler のフルインテール版に含まれています。
ディレクター	<i>SDF</i> ディレクター
データ	<i>Datos Meteorologicos</i>
アクター	<i>Datos Meteorologicos</i> , <i>RExpression</i> , <i>Display</i> , <i>ImageJ</i>
パラメーター	<p><i>Datos Meteorologicos</i>: Data Output Format= As Column Vector</p> <p><i>SDF</i> ディレクター: iterations = 1</p> <p><i>RExpression</i>: R function or script =</p> <pre> res <- lm(BARO ~ T_AIR) res plot(T_AIR, BARO) abline(res) </pre> <p><i>RExpression</i>: 入力ポート追加 'T_AIR' と 'BARO'</p>

線形回帰ワークフローは EarthGrid 上のデータの検索を行います。線形回帰を実行するワークフローがこれらのデータを使います。この例では、入力データが、*Datos Meteorologicos* アクター (2001年に La Hechicera ステーションで取られたデータ) の二つの出力ポート (気圧と気温のデータ列) から与えられます。

線形回帰ワークフローは4つのアクター (*Datos Meteorologicos* アクター、*RExpression* アクター、*ImageJ* アクター、*Display* アクター) と *SDF* ディレクターを使います。*RExpression* アクターは R のコマンドとスクリプトをワークフローに挿入します。*RExpression* アクターを使うと、R の強力なデータ操作と統計機能を簡単にワークフローに取り入れることができます。*RExpression* アクターを使うには、Kepler を実行するコンピュータに R インストールされていなくてはなりません。

注意: もしこのワークフローをうまく作成できなかったら、Kepler に付属の保存版ワークフローが次の場所にあります。kepler/demos/getting-started/05LinearRegression.xml

線形回帰ワークフローを作成するには、

1. コンポーネント・データアクセスエリア(コンポーネント・データアクセス エリアから「データ(Data)」タブを選ぶ。
2. 「Sources (ソース)」ボタンをクリックし、“KU Query Interface”と“KNB Metacat Authenticated Query Interface”のチェックマークをはずして検索範囲を制限します。Datos Meteorologicos は KNB Metacat (メタキャット) にあるので、検索対象とする EarthGrid 上のデータソースをこのノードのみに制限することができます。
3. Ok をクリックして検索ソースの変更を承認します。
4. 検索ボックスに Datos Meteorologicos とタイプして「Search (検索)」をクリックします。検索結果がかえって来るまでに 20 秒ほどかかるかもしれません。
5. 検索結果から Datos Meteorologicos アイコンをクリックします。Datos Meteorologicos アクターをワークフローキャンバスにドラッグします。

注意：このデータセットに関する詳しい説明を見るには、ワークフローキャンバス上の Datos Meteorologicos を右クリックし、「Get Metadata」を選択します(図24)。データ提供者の入力した情報量にもよりますが、沢山の貴重なメタデータを見ることができます。各フィールド属性の値の型や計測タイプなどはどのような統計モデルを実行するのが適切かを判断するための助けとなるでしょう。

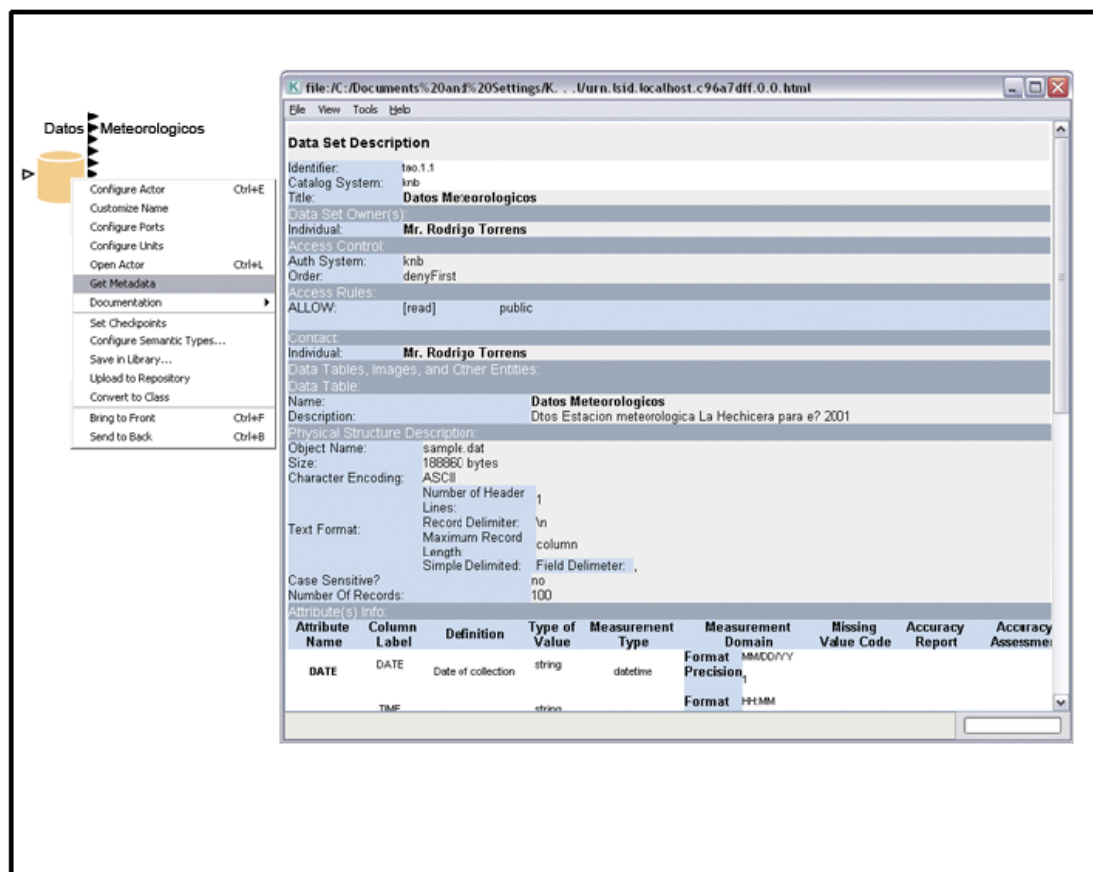


図 24: メタデータを閲覧する

6. *Datos Meteorologicos* アクターを右クリックし、「Configure Actor」を選択します。Data Output Format のそばのプルダウンメニューから「As Column Vector」を選び(図25)、Commit をクリックします。(*Datos Meteorologicos* アクターのデータ型は RExpression (R エクスプレッション) アクターの入力データの条件に合わせるため、“As Column Vector”に設定しなくてはなりません。)



図 25: Datos Meteorologicos を設定します。

注意: *Datos Meteorologicos* は、データ属性名に対応する出力ポートを持っています (例: BARO、T_AIR など)。適切なポートを見つけるには、各出力ポートの上にマウスを動かし、ポートのツールチップを確認してください。(図26)。

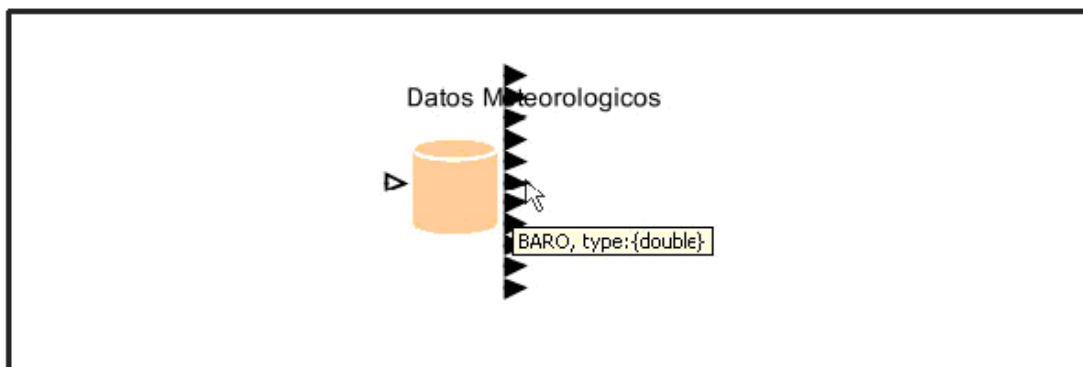


図 26: データポートを見つける。マウスを各ポートの上に移動させ、ツールチップを確認。

SDF ディレクターと残りのアクター(*RExpression*、*ImageJ*、*Display*)を作成してワークフローを完成させましょう。

7. *SDF* ディレクターを見つけてワークフローキャンバスにドラッグします。
8. *SDF* ディレクターを右クリックし、「Configure Director」を選択して *iterations* を 1 に変更して設定します。
9. 変更を有効にするため、「Commit」をクリックします。
10. *RExpression* アクターを見つけてワークフローキャンバスにドラッグします。*RExpression* アクターは“General Purpose”ファミリーに入っています。

初期設定では *RExpression* アクターは二つの出力ポートと 2+2 という R スクリプトが設定されています。 *RExpression* アクターを線形回帰ワークフローで使う前に、二つの入力ポート (T_AIR and BARO) を追加し、 *RExpression* スクリプトを修正しなくてはなりません。

11. *RExpression* アクターを右クリックし、「Configure Ports」を選択します。
12. “Configure ports” ダイアログボックスで、「Add」を二回クリックして新しいポートを二つ追加します。各ポートの隣の「Input」というチェックボックスをクリックして、これらの新しいポートを入力ポートに指定します。
13. 新しいポートの空白の Name 欄をダブルクリックし、新しい名前をつけます。一つには“T_AIR”、もう一つには“BARO”と名前をつけましょう。(図 27)。

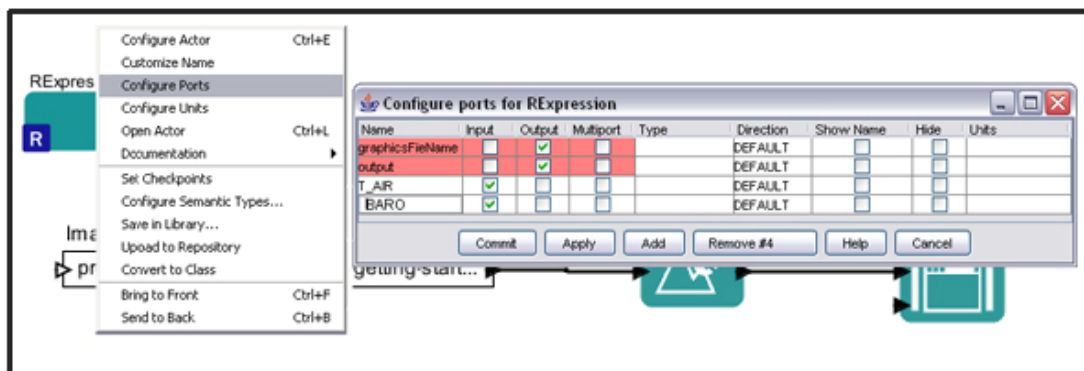


図 27: ポートを追加しカスタマイズする

14. R スクリプトを設定するために、 *RExpression* アクターを右クリックし、Configure Actor を選択します。“R function or script”ダイアログボックスの中で、R function or script の値を初期設定値から次の値に書き換えます。

```
res <- lm(BARO ~ T_AIR)
res
plot(T_AIR, BARO)
abline(res)
```

上の R 関数は、 *RExpression* アクターに気圧と気温の値を読み込んでそれらの値をプロットし、値を元に回帰直線を描くように指示を出します。変更を保存するために Commit をクリックします。

15. *Display* アクターをワークフローキャンバスへドラッグ・アンド・ドロップします。 *Display* アクターは“Components> Data Output > Workflow Output > Textual Output”というパスから見つけることができます。
16. *RExpression* アクターの下側の出力ポートを *Display* アクターの入力ポートに接続します。

17. *ImageJ* アクターをワークフローキャンバスへドラッグ・アンド・ドロップします。アクターは“Components > Data Output > Workflow Output > Graphical Output”というパスから見つけることができます。

*R*Expression アクターの上の出力ポートを *ImageJ* アクターのインポートポートに接続します。これでワークフローを実行する準備ができました。ワークフローの実行結果と画像出力は下のようになります(図28)。

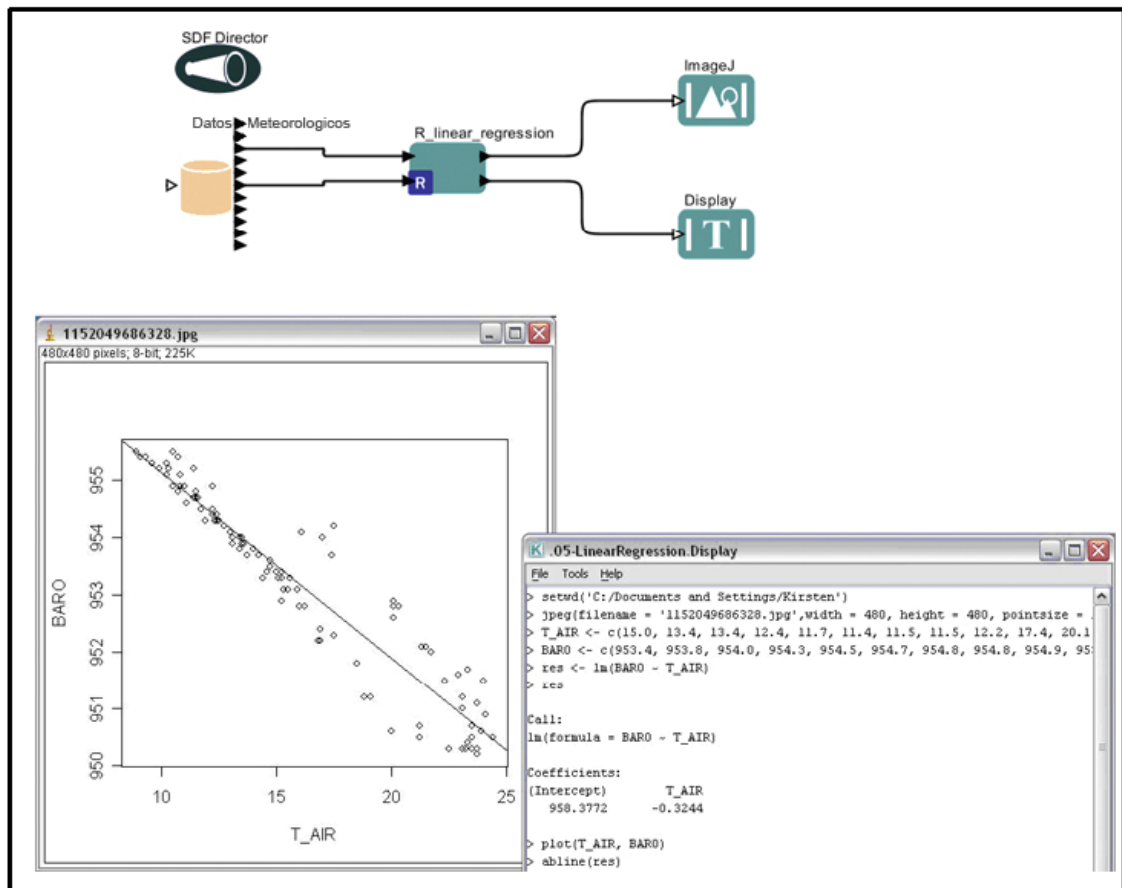


図 28: 線形回帰ワークフローとその出力

図28の左のウィンドウでは回帰直線の周りに気圧と気温の値をプロットした散布図を表示しています。グラフではこれら二変数の間に強い負の関係があることを示しています。つまり、気温が低ければ低いほど気圧が高くなるという関係です。右のウィンドウでは散布図で使われた気圧と気温のデータを表示します。またさらに、Y 軸の切片(線形回帰式 $y=mx+b$ において、気圧 958.38 と傾き -0.32)が表示されています。

ここまでできたら、他のデータタイプや他のデータセットを使ってワークフローを実行してみる事ができます。データを変えてみる場合は、セクション 7.2 のはじめのワークフロー表の仮定を満たすデータを使うことを忘れないでください。

7.3. サンプルワークフロー3-Web サービスとデータ変換

名前	Web サービスとデータ変換
ファイル名	06-WebServicesAndDataTransformation.xml
詳細記述	このワークフローは、特定の遺伝子アクセス番号に対応する塩基配列を検索するため、リモートのゲノクスデータベースサービスを利用します。この配列は適切な変換作業を経て3つの異なった形で表示されます。最初はネイティブフォーマット(XML)で、次はこのXMLフォーマットから抽出された配列の要素で、三番目はwebサイトで表示可能なHTMLドキュメントの形です。後の2つの操作は、一部の複雑な低レベル処理を隠したコンポジットアクターを使って実行されます。コンポジットアクターは、複雑になりそうな複数タスクを単一アクターで実行する、いふなればサブワークフローです。
仮定	Web Service(web サービス)アクターは、目的のweb サービス* <small>訳注</small> がRPC** <small>訳注</small> ベースで、プリミティブなXML型と配列を使っていると仮定します。
ディレクター	SDFディレクター
データ	データは初期入力の遺伝子アクセス番号がString Constant(文字列定数)アクターとして特定されていて、中間的な入力(XMLデータ)はリモートのゲノクスデータベースから検索したものです。
アクター	String Constant, Display, Sequence Getter Using XPath, HTML Generator Using XSLT, Web Services,
パラメーター	Web Services アクター: wsdlUrl = http://xml.nig.ac.jp/wsdl/DDBJ.wsdl methodName = getXMLEntry

*訳注 web サービス: ネットワークを介した協同的なコンピューター同士の相互利用を支援するソフトウェアデザイン("a software system designed to support interoperable machine-to-machine interaction over a network", World Wide Web Consortium)

**訳注 RPC: Remote Procedure Call, リモートホストのコンピュータープログラム(メソッド)を簡単に呼び出すことができ、結果をローカルコンピューターで利用できる仕組み

「web サービスとデータ変換」ワークフローはWeb Serviceアクターを使ってゲノクスデータベースへアクセスし、リモートのゲノクスデータベースサービスへ問い合わせして得られた、そこから塩基配列が返されます。戻り値の塩基配列の名前(つまり遺伝子アクセス番号)は、String ConstantアクターからWeb Serviceアクターに渡されます。Web Serviceアクターは適当なリモートサーバーに接続するように設定しなければなりません。一度設定されると、Web Serviceアクターはリモートサーバーから取得した塩基配列を出力し、3つの異なったテキスト表示するためのDisplayアクターを使って複数の形式で表示することができます: 一つはXMLで(サーバーがデフォルトで返す結果がこの形式です)、もう一つはこのXMLから抽出した配列の要素で、残りはwebサイトで表示可能なHTMLドキュメントです。

リレーションはWeb Serviceアクターによって出力されるデータを分岐させるのに使われ、データを必要とする全てのコンポーネントの間で共有されることとなります。

このワークフローは2つのコンポジットアクターを使用します:*Sequence Getter Using XPath*(Xパス^{*訳注}を使用した配列取得)と *HTML Generator Using XSLT*(XSLT^{**訳注}を利用したHTML生成)で、それぞれXMLデータを返すのと配列要素とHTMLファイルへ変換するのに使われます。このコンポジットアクターは既存のKeplerアクターを使って、このワークフローの中で使用するために生成されたものです。*Sequence Getter Using XPath*と *HTML Generator Using XSLT*という名前のコンポーネントは、「コンポーネント」タブには現れません。コンポジットアクターの内部を見たい場合は、ワークフローキャンバスでこのアクターを右クリックし、メニューから「Open Actor(アクターを開く)」を選択して下さい。コンポジットアクターは新しいアプリケーションウィンドウで開きます。図32はその例です。

*訳注 XML Path Language, XML ドキュメント内の場所を特定する言語

**訳注 XML Stylesheet Language Transformations, Xパスを利用して、あるXML文書を他のXML文書や異なった形式の文書へ変換するための言語

加えて、このワークフローは4番目の *Display* アクターを使い、リモートサーバーが返したエラーを表示します(例えば、サーバーダウンや不正な入力)。

この「webサービスとデータ変換」ワークフローを作成するには：

1. 新しいワークフローキャンバスを開きます。
2. *SDF* ディレクターをワークフローキャンバスへドラッグ・アンド・ドロップします。パラメーター *iterations*(反復数)を1にします。
3. *String Constant*(文字列定数)アクターをワークフローキャンバスへドラッグ・アンド・ドロップします。
4. *String Constant* アクターを右クリックし「Configure Actor(アクターの設定)」を選択します。Value(値)フィールドにAA045112を入力し、「Commit(承認)」ボタンを押します。
5. *String Constant* アクターの名前を変えるため、これを右クリックし、「Customize Name(名前の変更)」を選びます。新しい名前(例えば Gene Accession Number)を New name(新名)フィールドへ入力し、「Commit」ボタンを押します(図29)。

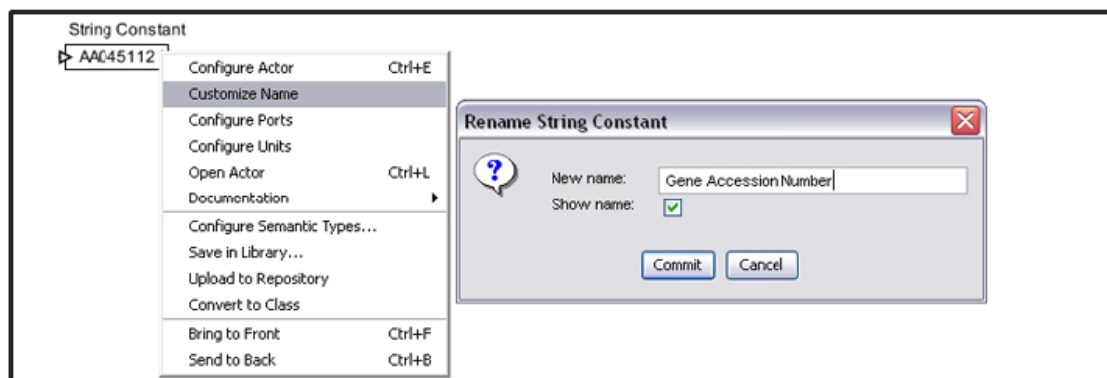


図 29: アクターの名前を変更

6. *Web Service*(webサービス)アクターをワークフローキャンバスへドラッグ・アンド・ドロップします。*String Constant* アクターの下へ置いて下さい。デフォルト

では、*Web Service* アクターは実行時エラーを出力するための出力ポートを一つ持ちます、また *web* サービスの URL(`wsdUrl` パラメーター)とそのメソッド^{*訳注}(`methodName` パラメーター)を設定しなければなりません。これらの情報によって一度アクターが設定されると、*web* サービスによって要求される適切な入力ポートが生成されます。

7. *web* サービスにアクセスするために必要なパラメーターを設定するため、*Web Service* アクターを右クリックし、「Configure Actor(アクターの設定)」を選択します(図 30)。`wsdUrl` に `http://xml.nig.ac.jp/wsd/DDBJ.wsdl` を入力します^{**訳注}。`methodName` へ `getXMLEntry` を入力します。*Web Service* アクターのポートは自動的に更新されるはずですが、アクターを右クリックして *Configure Ports*(ポート設定)ダイアログボックスから「*direction*(アクターでのポートの位置)」のプルダウンメニューを選ぶことで、より便利な場所へポートを移動することができます^{***訳注}。

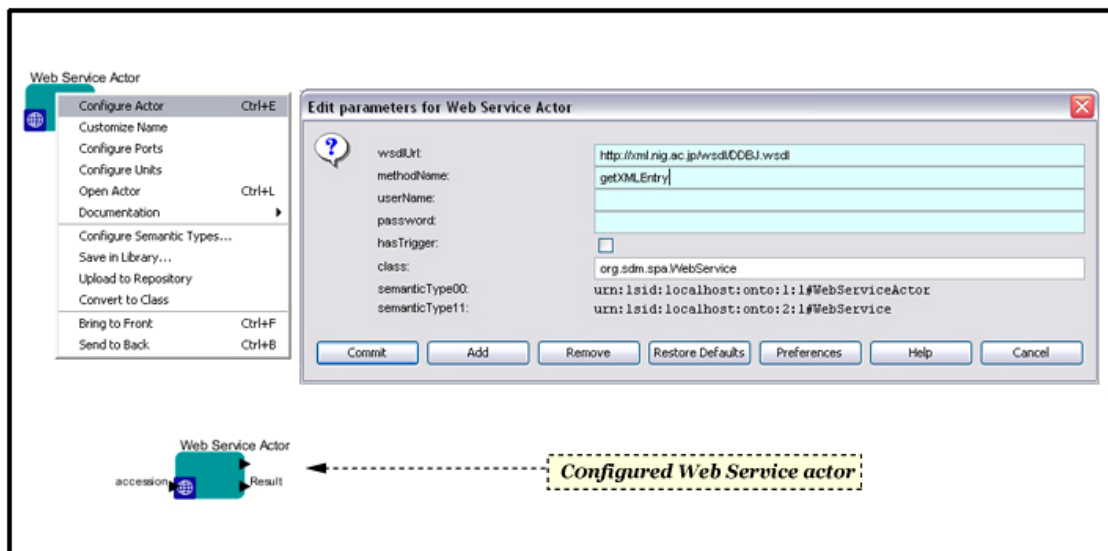


図 30 *web* サービス(*Web Service*)アクターの設定

^{*訳注} ここでのメソッドは *web* サービスを利用する際の、*Web* サービス記述言語 WSDL で書かれた XML 文書の名前で `getXMLEntry` や `searchParam` など

^{**訳注} `wsdUrl` だけ一度入力して `commit` すると、`methodName` はプルダウンメニューで選ぶことができるようになる

^{***訳注} 後半の記述はポートの場所を任意に動かせるように書いているが、アクターの左右上下(東西南北)しか選べない

8. *Gene Accession Numbers* という名前を付けた *String Constant* アクターの出力を、*Web Service* アクターの入力につなぎます。
9. *Display*(表示)アクターを 4 つワークフローキャンパスへドラッグ・アンド・ドロップします。
10. 一つの *Display* アクターを *Web Service* アクターの右下へ持ってきます。この *Display* アクターを右クリックして、*Errors Sink* という名前に変更します。
11. *Web Service* アクターの上の方の出力ポートを *Errors Sink* の *Display* アクターへ接続します。

- 2番目の *Display* アクターを web サービスアクターの右の少し上に持ってきます。アクターを右クリックして、"XML Entry Display"と名前を変更します。

「web サービスとデータ変換」ワークフローは、このワークフローのために特別に設計された2つのコンポーネントアクターを使います。これらのカスタマイズされたアクターはコンポーネントライブラリの中で利用可能とはならず、むしろそれらを再び作る必要があります。既存のワークフローからそれらをコピー・ペーストすることで、これを利用します。

- Kepler をインストールしたディレクトリの "/demos/getting-started" から web サービスとデータ変換とを行うワークフロー"(06-WebServicesAndDataTransformation.xml"を開きます。ワークフローは新しいアプリケーションウィンドウで開きます。 *Sequence Getter Using XPath*(Xパスを使用した配列取得)コンポジットアクターを左クリックして選択します。
- 「Edit(編集)」メニューから「Copy(コピー)」を選びます(あるいは、キーボードショートカット[Ctrl]+[C]を使います)
- 現在作成中のワークフローへ戻り、コピーした *Sequence Getter Using XPath* コンポジットアクターを編集メニューの「Paste(貼り付け)」を使うか(図 31)キーボードショートカットの[Ctrl]+[V]を使って貼り付けます。

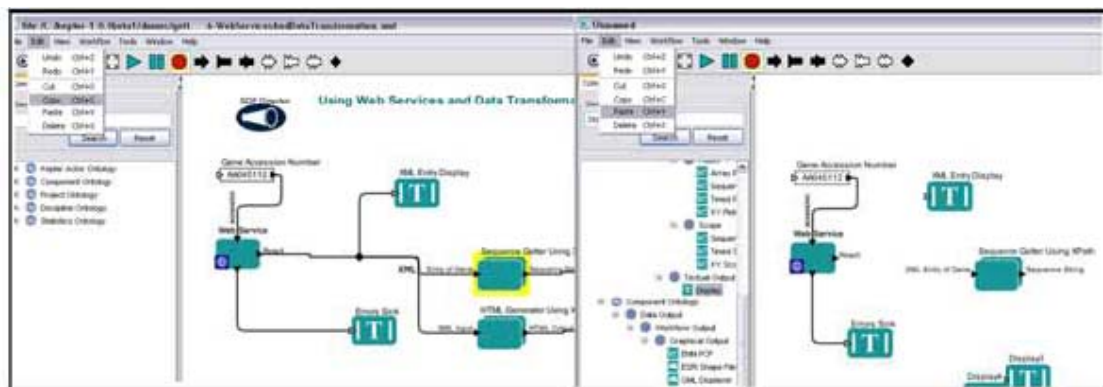


図 31: ワークフロー間でのアクターのコピーと貼り付け

- オリジナルの「web サービスとデータ変換」ワークフロー (WebServicesAndDataTransformation) から *HTML Generator Using XSLT*(XSLT を利用した HTML 生成)コンポジットアクターを作成中のワークフローへコピーと張り付けをします

注意: コンポジットアクターの内部を見たい場合は、アクターを右クリックし、メニューの「Open Actor(アクターを開く)」を選択します。コンポジットアクターが新しいアプリケーションウィンドウの中に開きます(図 32)。コンポジットアクターは複雑なタスクの組み合わせを実行するサブワークフローと考えることができます。

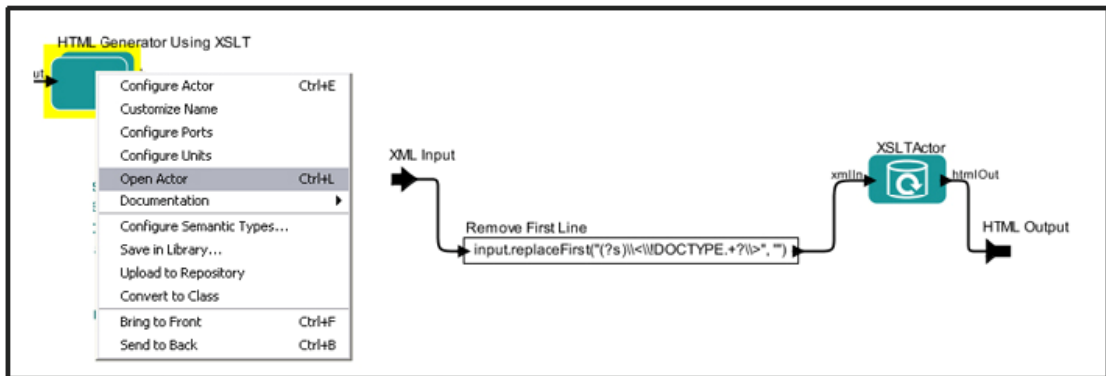


図 32 HTML Generator Using XSLT コンポジットアクターの内部

Web Service アクターの出力は3つのアクターに要求されているので、アクター接続をする前に、出力を分岐させるためにリレーションを追加する必要があります。

17. ツールバーの右端にある Relation(リレーション)アイコンをクリックしてリレーションを追加します。リレーション(黒いひし形アイコン)はワークフローキャンパスの中央付近に現れます(図33)。キーボードショートカットの[Ctrl]+[左クリック]でもリレーションを追加できます。

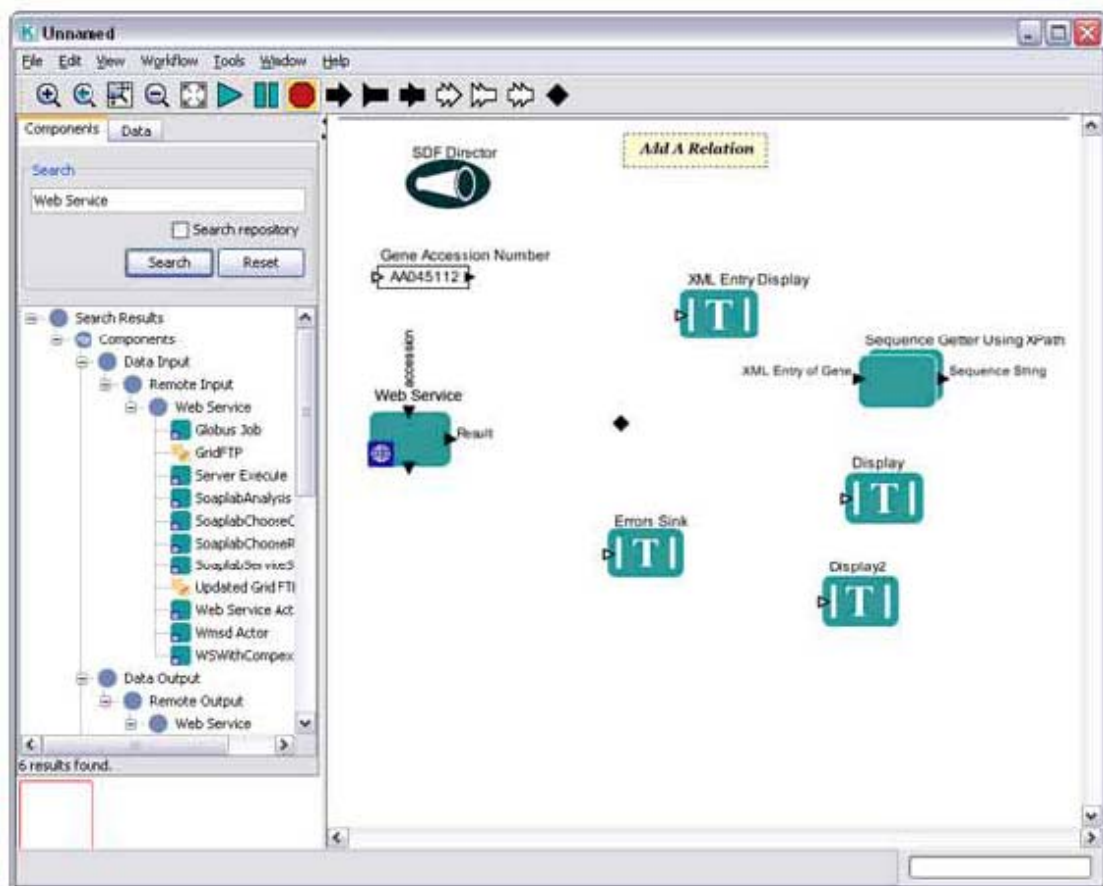


図 33 Relation(リレーション)の追加

18. リレーションアイコンを *Web Service* アクターと *Sequence Getter using XPath* アクターの間に持ってきます。
19. “XML Entry Display” *Display* アクターと *Web Service* アクターを接続します。この *Display* アクターの入力ポートをクリックしリレーションアイコンへドラッグ・アンド・ドロップします。
20. *HTML Generator Using XSLT* アクターと *Sequence Getter Using XPath* アクターも同様にリレーションアイコンと接続します。
21. 3 番目の *Display* アクターを「Sequence String Display」と名付け、*Sequence Getter using XPath* アクターの右に置きます
22. *Sequence Getter using XPath* アクターの出力を”Sequence String Display”アクターの入力に接続します
23. 4 番目の *Display* アクターを”HTML Display”と名付け、*HTML Generator Using XSLT* アクターの右に置きます。
24. *HTML Generator Using XSLT* アクターの出力を”HTML Display”アクターの入力へ接続します。

これでワークフローの実行準備が整いました。ワークフローの結果と出力は次のようになります(図 34)

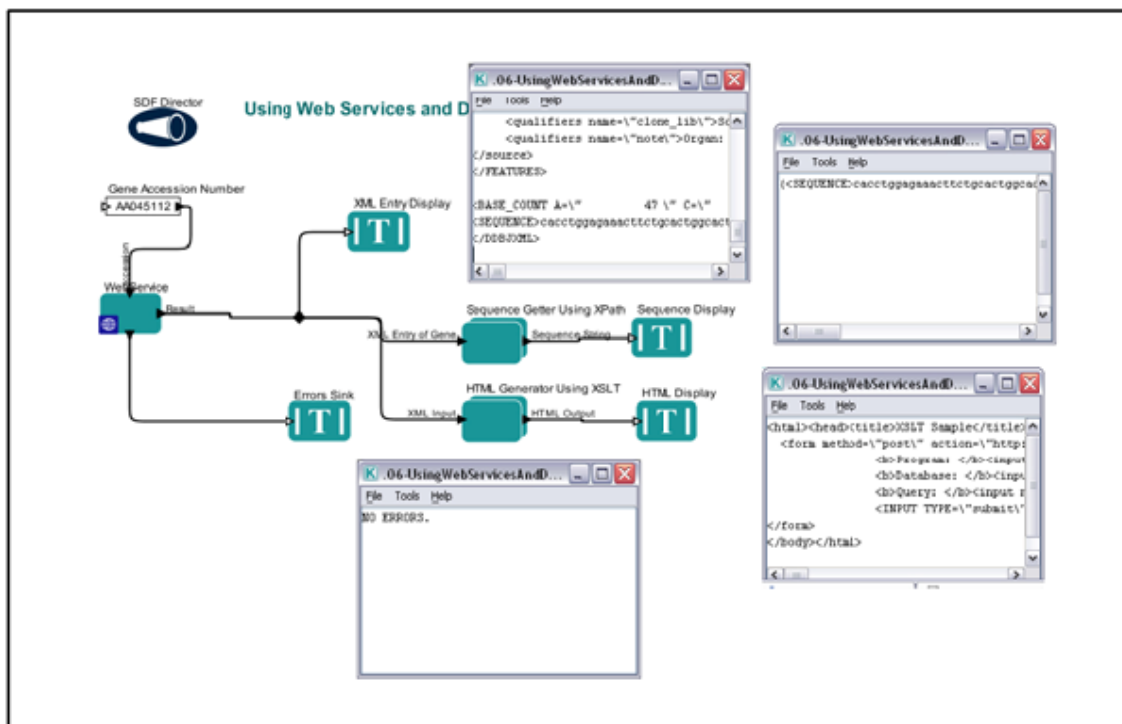


図 34 web サービスワークフロー

注意： 作成したワークフローに注釈をいれるには、Annotation(注釈)アクターをワークフローキャンバスへドラッグ・アンド・ドロップします。ダブルクリックすることでデフォルトのテキスト“Double click to edit”(ダブルクリックして編集)を変更することができます。

7.4. サンプルワークフロー4-Kepler からの外部アプリケーションの 実行 (External Execution(外部実行)アクター)

External Execution(外部実行)アクターは Kepler の中から外部アプリケーションを起動するために使います。このアクターはパラメーターを受け渡し、実行結果の値を返します。これは下流のアクターが使ったり表示したりすることができます。外部実行アクターを使用するためには、起動させるアプリケーションはローカルコンピュータで実行され、場合によっては、適切に設定されていなければなりません。このセクションでは、外部実行アクターを使うワークフロー例を見て行きます。

名前	外部実行ワークフロー
ファイル名	07-CommandLine_1.xml
詳細説明	"07-CommandLine_1.xml"ワークフローは Kepler の External Execution (外部実行) アクターを利用して、Kepler のパッケージの中にある Hello World Java アプリケーションを実行します。このアクターはアプリケーションの結果を出力し、表示アクターで表示することができます。
仮定	Hello World Java アプリケーションがローカルマシンの Kepler をインストールしたディレクトリの/demos/getting-started ディレクトリにあることを想定しています。
ディレクター	SDF ディレクター
データ	2つの <i>Constant</i> (定数)アクターからデータが提供されます
アクター	<i>Constant</i> , <i>External Execution</i> , <i>Display</i>
パラメーター	<i>External Execution</i> アクター: directory = \$WorkingDir waitForProcess が選択(チェックされている状態)

Command Line 1 ワークフローは Kepler の *External Execution*(外部実行)アクターを利用して、Kepler のパッケージの中にある HelloWorld Java アプリケーションを実行します。HelloWorld Java アプリケーションは単純な Java プログラムで、"Hello"というテキストに変数(たいていはユーザー名、そのデフォルトは"Kepler_User"という文字列)という文字列が出力されます。外部実行アクターは HelloWorld Java アプリケーションの実行が終わるまで待機し、アプリケーションの出力を結果として返し、これは表示アクターで表示できます。

外部実行アクターの *directory*(ディレクトリ)パラメーターは HelloWorld アプリケーションの場所を設定します。他の全てのパラメーターはデフォルト設定のままです。

Command Line 1 ワークフローを作成するには:

1. *SDF* ディレクターをワークフローキャンパスヘドラッグ・アンド・ドロップします。パラメーター *iterations*(反復数)を 1 にします。

2. **Constant(定数)**アクターをワークフローキャンパスへドラッグ・アンド・ドロップします。このアクターの名前を **CommandLine** とします。アクターを命名するには、アクターを右クリックしてドロップダウンメニューの「**Customize Name(名前を変更)**」を選びます。新しい名前を **New name(新しい名前)**フィールドへ入力し、「**Commit(承認)**」ボタンを押します。
3. 「**CommandLine**」アクターの中央をダブルクリックして **Edit Parameters(パラメーター編集)**ダイアログボックスを表示させます。パラメーターvalue を"**java -cp ./ HelloWorld Kepler_User**"とします(引用符"も入力して下さい)。最後の **Kepler_User** というのはコマンドラインを通して伝えられる、Java プログラムへの引数です。この値は希望するものへと変えることができます(たとえば、**Katie** や **Bob**)。この文字列全体を囲んでいる引用符"はこれが文字列であることを明示するために必要です。「**Commit 承認**」ボタンを押します。
4. **Component** タブで、**Component library** の中から **Parameter** を検索して下さい、そして **Workflow** のツリーから **Parameter(青い丸)**をワークフローキャンパスへドラッグ・アンド・ドロップします。このアクターを右クリックし、ドロップダウンメニューの「**Customize Name(名前を変更)**」を選びます。**WorkingDir** と命名し、「**Commit(承認)**」ボタンを押します。このアクターを今度はダブルクリックし、**Edit Parameters(パラメーターを編集)**ダイアログボックスを表示させ、**property ("KEPLER")+"/demos/getting-started"**と入力します。これはつまり **working directory(作業ディレクトリ)**の場所です。ここでは **Kepler** をインストールしたディレクトリの下の **"/demos/getting-started"**ディレクトリが作業ディレクトリとなります。
5. 「**コンポーネント**」タブで「**External Execution**」を検索し、このアクターをワークフローキャンパスへドラッグ・アンド・ドロップします。アクターをダブルクリックして、パラメーター**directory(ディレクトリ)**を**\$WorkingDir** とします(つまりパラメーター**WorkingDir** は4.で追加したワークフローパラメーターで設定されることになります)。(図35)



図 35 このワークフローで外部実行アクターを使うためにパラメーターdirectory(ディレクトリ)を設定

6. 「CommandLine」 *Constant* アクターの出力ポートを *External Execution* アクターの *command*(コマンド)入力ポートへ接続します。 *command* ポートはアクター左側の真ん中です(見づらいときはワークフローキャンバスをズームします)
7. *Display*(表示)アクターをワークフローキャンバスへドラッグ・アンド・ドロップし、 *External Execution* アクターの *output*(出力)ポートを、このアクターの *input*(入力)ポートへ接続します。
8. ワークフロー実行の準備が整いました。このワークフロートデフォルトの出力は 図36 に示しました。

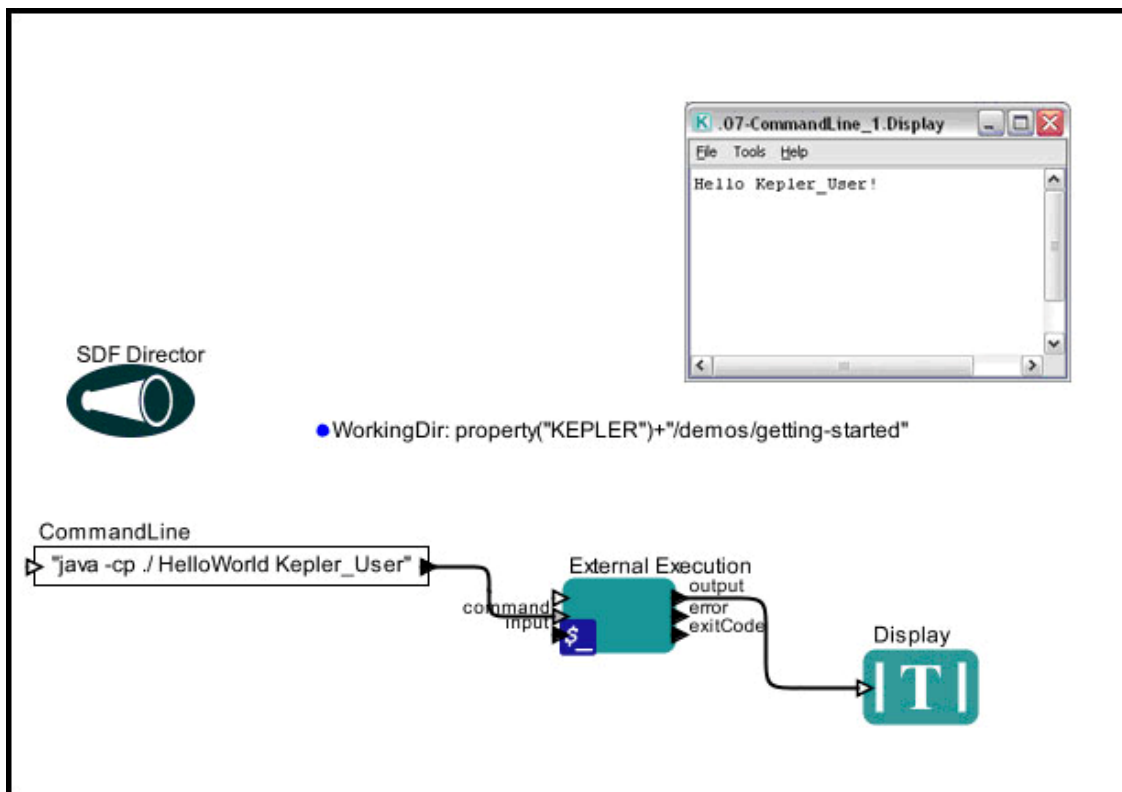


図 36 Command Line 1 ワークフローとデフォルトの出力

8. 付録

8.1. Ptolemy II(トレミー2)-Kepler の基礎

Ptolemy II(トレミー2)^{*訳注}は不均質さと協調性を可能にするモデリングとデザインのためのソフトウェア フレームワークであり、Java ベースのコンポーネントフレームワークに Vergil(バージル)^{**訳注}というグラフィカルインターフェイスを使っています。Ptolemy II はカリフォルニア大バークレー校の Ptolemy プロジェクトの成果であり、このプロジェクトは「明確な計算モデルの利用によるコンポーネントの相互作用の決定」がゴールでした。

プロジェクト web サイトの説明によれば、Ptolemy II はいくつかのドメインを含み、各

ドメインはある計算モデルを実現します。コンポーネントライブラリと、グラフィックス、数学、プロット、データなどのサポートパッケージを持ちます。Ptolemy II のより詳しい情報は <http://ptolemy.eecs.berkeley.edu/index.html> を参照して下さい。

*訳注 名前の由来は、2世紀の天文学者、数学者、地理学者のプトレマイオスより

**訳注 Vergil(バージル, ウェルギリウスとも) 古代ローマの詩人

科学ワークフローをもともと意図していないにも関わらず、Ptolemy II はデータフロー的—これは科学ワークフローの大変重要な特徴です—モデル開発支援を提供します。Ptolemy II はオープンソースで、様々な計算モデルを含んだモデルデザインと実行の成熟したプラットフォームであり、充実したドキュメントと容易な拡張性を提供しているので、Kepler の基礎として選択しました。

8.2. アクターリファレンス

アクターとディレクターのドキュメントは Actor Reference(アクターリファレンス)にあります。また、Kepler インターフェイスの中でも見ることができます。ドキュメントを見るには

1. アクターやディレクターを右クリックします。
2. 「Documentation(ドキュメント)」を選びます。
3. 次に「Display(表示)」を選択します(図 37)。

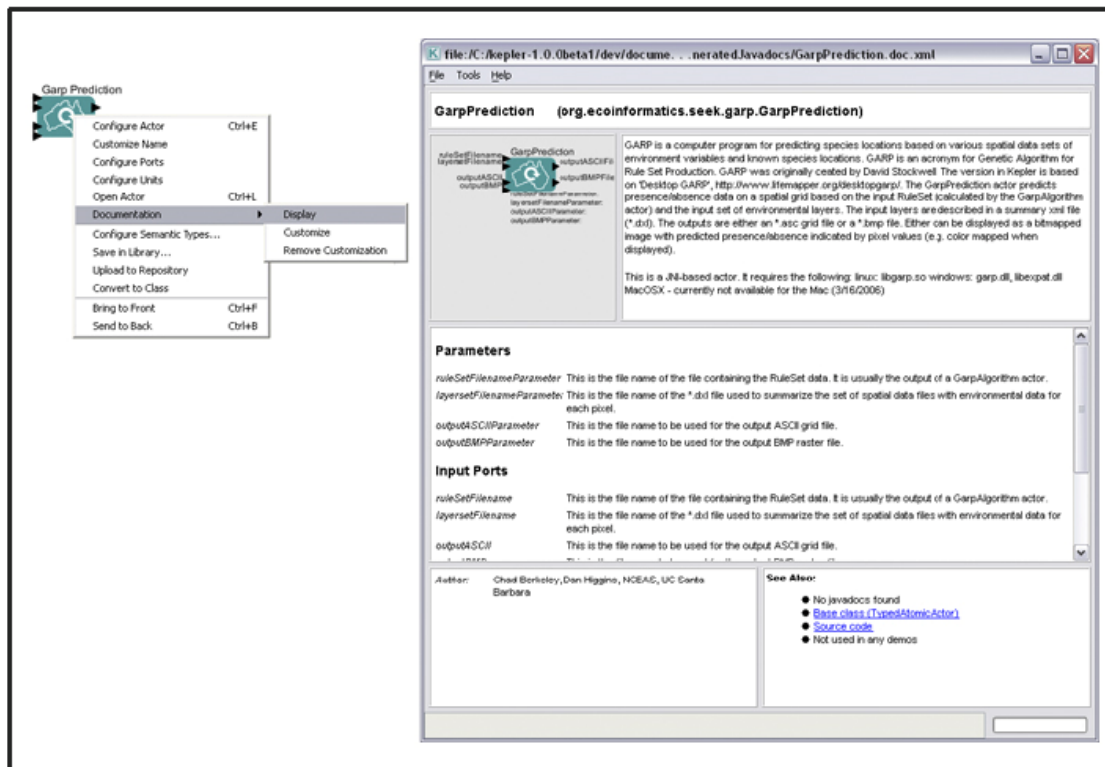


図 37 アクターのドキュメント