

TeX Live パッケージと Neo4j

Norbert Preining
norbert@preining.info

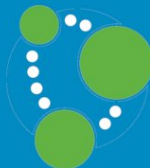


www.accelia.net www.tug.org/texlive



Neo4j ユーザー勉強会 #18

2018 年 9 月 26 日



概要

- ▶ 自己紹介
- ▶ T_EX Live とは何か？
- ▶ T_EX Live のパッケージ組織
- ▶ Neo4j での表現
- ▶ パッケージグラフでの模索
- ▶ グラフ・アルゴリズム

自己紹介

数学と情報学

We also consider the case of \rightarrow -conversion in detail. Consider an \rightarrow -conversion redex occurring in a hyper natural deduction (R, \preceq) of the form

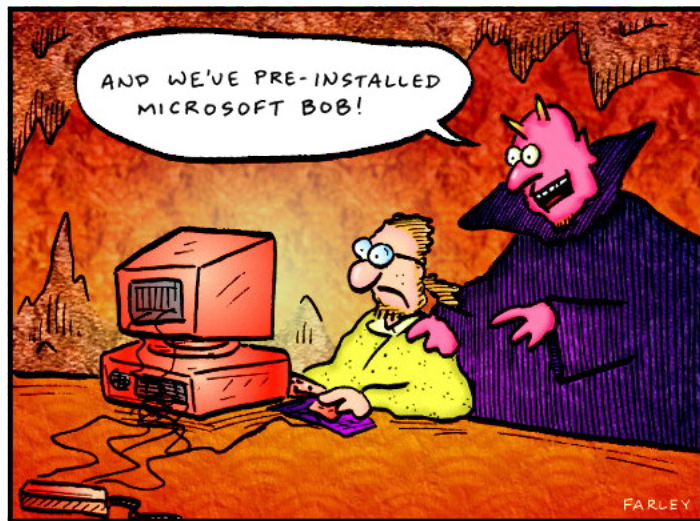
$$\frac{\frac{\frac{\sigma_1 \vdots}{t} \frac{B}{A \rightarrow B} \quad \frac{\sigma_2 \vdots}{s} A}{\frac{1}{x: \rightarrow-i} \frac{y: \rightarrow-e}{B}}}{B}$$

If no assumption of σ_1 is discharged at \rightarrow -i, then this redex converts to

$$\frac{\sigma_1 \vdots}{B}$$

20+ 年間：オーストリア（ウィーン工科大学）、イタリア（シエナ大学）、日本（北陸先端科学技術大学院大学）

Debian Developer



Copyright © 1995 David Farley, d-farley@tezcat.com
Distributed by United Feature Syndicate
<http://www.unitedmedia.com>

This cartoon is made available on the Internet for personal viewing only.

だいたい $\text{T}_\text{E}\text{X}$ 関係のパッケージ (プラス Calibre, mu, ...)

T_EX Live 2018



T_EX Live (メインインフラ開発者)、日本語サポート
(texjp.org)

国際山岳ガイド



岩登り、山スキー、アイスクライミング、ハイキング、沢登り、...

アクセリア株式会社・研究開発部

CDN、ウェブサービス、セキュリティーなどのサービス¹



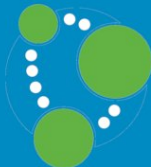
セキュリティー、機械学習、ソフトウェア検証、ブロックチェーン等

¹<http://www.accelia.net>

T_EX Live とは何か？

T_EX Live

- ▶ T_EX とそのまわりのプログラム : pdftex, luatex, ptex, uptex, dvips, dvipdfmx, mf, xetex, detex, ht, patgen, ...

T_EX Live

- ▶ T_EX とそのまわりのプログラム : pdftex, luatex, ptex, uptex, dvips, dvipdfmx, mf, xetex, detex, ht, patgen, ...
- ▶ L^AT_EX macro and font package:
memoir, beamer, lm, dejavu, fira, ipaex, libertine, ...

T_EX Live

- ▶ T_EX とそのまわりのプログラム : pdftex, luatex, ptex, uptex, dvips, dvi_{pdf}mx, mf, xetex, detex, ht, patgen, ...
- ▶ L^AT_EX macro and font package: memoir, beamer, lm, dejavu, fira, ipaex, libertine, ...
- ▶ T_EX Live Manager: online update、設定、検索など

フィーチャー

- ▶ ‘完成’ – 全ての CTAN にある OSS パッケージを含む
- ▶ マルチプラットフォーム、同じシステム
Windows == Unix (*cum grano salis*)
- ▶ 毎日の更新
- ▶ DFSG フリー
- ▶ 様々なインストール方法 (DVD, network, mirror of CTAN, svn checkout, 他のインストール)
- ▶ インストーラー: テキストと GUI モード
- ▶ 翻訳された (日本語が含まれてる)

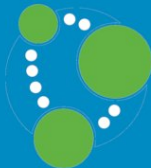
TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group



TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX



TLの歴史

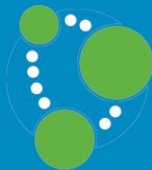
- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX
- ▶ 1996 first edition, Sebastian Rahtz





TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX
- ▶ 1996 first edition, Sebastian Rahtz
- ▶ 2000 5th edition, non-free software removed



TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX
- ▶ 1996 first edition, Sebastian Rahtz
- ▶ 2000 5th edition, non-free software removed
- ▶ 2002 7th edition: Mac OS X support



TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX
- ▶ 1996 first edition, Sebastian Rahtz
- ▶ 2000 5th edition, non-free software removed
- ▶ 2002 7th edition: Mac OS X support
- ▶ 2005 addition of the -sys scripts

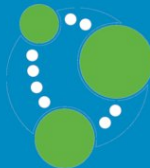
TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4All group
- ▶ 1995 Unix-based TDS CD based on te_EX
- ▶ 1996 first edition, Sebastian Rahtz
- ▶ 2000 5th edition, non-free software r
- ▶ 2002 7th edition: Mac OS X support
- ▶ 2005 addition of the -sys scripts
- ▶ 2006-09 XeT_EX addition, end of te_EX development, T_EXworks addition, Karl Berry



TL の歴史

- ▶ late 1993 Dutch T_EX Users Group, 4AllT_EX CD, tds working group
- ▶ 1995 Unix-based TDS CD based on teT_EX
- ▶ 1996 first edition, Sebastian Rahtz
- ▶ 2000 5th edition, non-free software removed
- ▶ 2002 7th edition: Mac OS X support
- ▶ 2005 addition of the -sys scripts
- ▶ 2006-09 XeT_EX addition, end of teT_EX development, T_EXworks addition, Karl Berry
- ▶ 2010- incremental inclusion of Japanese T_EX support

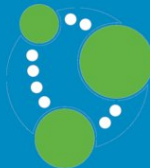


インストーラーの設定

▶ arch/os システム

arch-os を選択してください

- ☐ Cygwin on Intel x86
- ☐ Cygwin on x86_64
- ☐ MacOSX current (10.10-) on x86_64
- ☐ MacOSX legacy (10.6-10.9) on x86_64
- ☐ FreeBSD on x86_64
- ☐ FreeBSD on Intel x86
- ☐ GNU/Linux on ARM64
- ☐ GNU/Linux on ARMhf
- ☐ GNU/Linux on Intel x86
- ☒ GNU/Linux on x86_64
- ☐ GNU/Linux on x86_64 with musl
- ☐ NetBSD on x86_64
- ☐ NetBSD on Intel x86
- ☐ Solaris on Intel x86
- ☐ Solaris on Sparc
- ☐ Solaris on x86_64
- ☐ Windows



インストーラーの設定

▶ arch/os システム

▶ schemes

選択したスキーム

- ☒ 完全スキーム（すべて）
- ☐ 中規模スキーム（小規模スキームにパッケージと言語を追加）
- ☐ 小規模スキーム（基本スキーム + xetex · metapost · 少々と言語）
- ☐ 基本スキーム（plainおよびlatex）
- ☐ 最小限スキーム（plainのみ）
- ☐ ConTeXtスキーム
- ☐ GUST TeX Liveスキーム
- ☐ infrastructure-only scheme (no TeX at all)
- ☐ teTeXスキーム（中規模スキームよりも多いが完全スキームには及ばない）
- ☐ コレクションの任意選択

インストーラーの設定

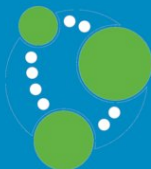
▶ arch/os システム

▶ schemes

▶ collections



T_EX Live のパッケージ組織



3つのレベル

- ▶ Package : 基本のユニット、CTAN のパッケージはだいたい T_EX Live の Package になる
beamer, memoir, pgf, ...
- ▶ Collection : 関係があるパッケージの集合
collection-fontsrecommended, collection-mathscience,
...
- ▶ Schema : トップレベル、collection と package を含む
scheme-basic, scheme-standard, scheme-full, ...

「含む」は depends という関係がある

TeX Live データベース : Package

```
name 12many
category Package
revision 15878
catalogue one2many
shortdesc Generalising mathematical index sets
longdesc In the discrete branches of mathematics and the
..
docfiles size=98
  texmf-dist/doc/latex/12many/12many.pdf details="Package
  texmf-dist/doc/latex/12many/README details="Readme"
srcfiles size=6
  texmf-dist/source/latex/12many/12many.dtx
  texmf-dist/source/latex/12many/12many.ins
runfiles size=1
  texmf-dist/tex/latex/12many/12many.sty
..
```

TeX Live データベース : Collection

```
name collection-langjapanese
category Collection
revision 48752
shortdesc Japanese
longdesc Support for Japanese; additional packages in
longdesc collection-langcjk.
depend collection-langcjk
depend ascmac
depend babel-japanese
depend bxbase
depend bxcjkatype
depend bxjalipsum
depend bxjaprnind
depend bxjsc1s
depend bxorigcapt
..
```

TeX Live データベース : Schema

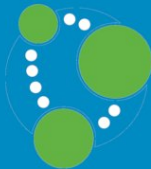
```
name scheme-medium
category Scheme
revision 44177
shortdesc medium scheme (small + more packages and langu
longdesc This is the medium TeX Live collection: it cont
longdesc LaTeX, many recommended packages, and support f
longdesc languages.
depend collection-basic
depend collection-binextra
depend collection-context
depend collection-fontsrecommended
depend collection-fontutils
depend collection-langczechslovak
depend collection-langenglish
depend collection-langeuropean
..
```

数・数・数

現在の T_EX Live に含まれてるのは :

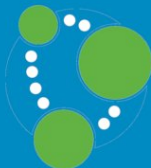
- ▶ 9 Scheme
- ▶ 41 Collection
- ▶ 6718 Package
- ▶ 181839 Files
- ▶ 17 architecture-os combinations

Neo4j ^



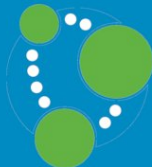
T_EX Live Database から Neo4j へ

- ▶ 簡単な Perl スクリプト
- ▶ T_EX Live の Perl Modules を使う
- ▶ CSV にエクスポート
- ▶ neo4j-import でインポート



CSV: node-Package.csv

```
uuid:ID,name
idde511a2208a14810b27f36ddbb0065a8,mkpic
ida7ce0776c7f64cefa2e2a74bd94268a1,mpcolornames
idc63eeda473104ebee21e5844dd7c5660,gfnotation
id3cfc0196bf5c4c1baeffabca67b9a544,turkmen
id3b602a6c491544deb2edd510c1fe2c20,cascadilla
...
```



CSV: node-Collection.csv

uuid:ID,name

id96c30fc211be44219457852dc835fd2a,collection-langczechs

id1788d667112443b9fe1eba27628498c,collection-langcyril1

id186afd7ec58e4285a09dbfe7650f6542,collection-langjapane

id100a6c30d11d443191ba4abab9be9d37,collection-wintools

id00b2ac0c815b48eb866029f530c525ad,collection-pictures

...



CSV: edge-depends.csv

```
:START_ID, :END_ID
```

```
idc9215c5d96bf4cb587785d8c0d50a6e0, id120a994a30204e40bf1
```

```
idc9215c5d96bf4cb587785d8c0d50a6e0, idf989960c4f494e1a90f
```

```
id96c30fc211be44219457852dc835fd2a, id2883c1628d4544cc8e4
```

```
id96c30fc211be44219457852dc835fd2a, id10536fb246e9471d86e
```

```
id96c30fc211be44219457852dc835fd2a, idc3e0c0d5add747e39ff
```

```
...
```

CSV: edge-contains.csv

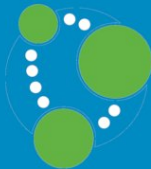
どのファイルはどのパッケージに含まれてる情報

```
:START_ID,type,:END_ID
id0ecf144c7bfc4c7a8037981121810d8d,doc,id82f6f55fe13d4e7
idf4166223dcdb446b95ac9d9e551ea562,run,ideba9104946c64d4
id6d704e156bed47d39f41ebd1ce447dae,run,id6616c86a71854f7
id6e0b59327a5942b49aa670f815f4a579,run,idee5780bcadb04ab
id469eb48326ce46e7a213d52fddfcc91a,doc,ide535515a72314cb
...
```

```
$ ls
edge-contains.csv      node-ConTeXt.csv      node-Scheme.csv
edge-depends.csv       node-Files.csv        node-TLCore.csv
node-Collection.csv    node-Package.csv
$ neo4j-import --into ../graphdb \
--nodes:Collection node-Collection.csv \
--nodes:ConTeXt node-ConTeXt.csv \
--nodes:Files node-Files.csv \
--nodes:Package node-Package.csv \
--nodes:Scheme node-Scheme.csv \
--nodes:TLCore node-TLCore.csv \
--relationships:contains edge-contains.csv \
--relationships:depends edge-depends.csv
...
IMPORT DONE in 4s 199ms.
Imported:
  168140 nodes
  168803 relationships
  336280 properties
Peak memory usage: 1.03 GB
```

パッケージグラフでの模索

簡単な Cypher での模索



簡単な Cypher での模索

全ての Schema

```
match (s:Scheme) return s;
```



簡単な Cypher での模索

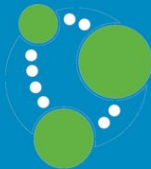
全ての Schema

```
match (s:Scheme) return s;
```

Collection 以外の Schema の依存

```
match p = (s:Scheme) -[:depends]-> (q)
  where NOT 'Collection' IN LABELS(q)
  return p;
```

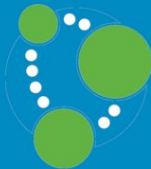
一貫性の確認



一貫性の確認

複数の Collection に入ってるパッケージがない

```
match (c1:Collection) -[:depends]-> (p)
      <-[:depends]- (c2:Collection) return c1, c2, p;
```



一貫性の確認

複数の Collection に入ってるパッケージがない

```
match (c1:Collection) -[:depends]-> (p)
      <-[:depends]- (c2:Collection) return c1, c2, p;
```

サイクルがない

```
match p = (n)-[:depends*]-> (n) return p;
```

一貫性の確認

複数の Collection に入ってるパッケージがない

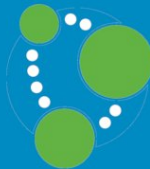
```
match (c1:Collection) -[:depends]-> (p)
  <-[:depends]- (c2:Collection) return c1, c2, p;
```

サイクルがない

```
match p = (n)-[:depends*]-> (n) return p;
```

複数のパッケージに入ってるファイルがない

```
match (p1) -[:contains]-> (f)
  <-[:contains]- (p2) return p1, p2, f;
```

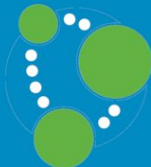


ファイルタイプの検索

1つのパッケージの全てのドキュメンテーションファイルの検索

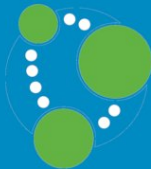
```
match (p) -[:contains {type:'doc'}]-> (f)
  where p.name = "tlcockpit"
  return p,f;
```

グラフ・アルゴリズム



グラフ・アルゴリズムライブラリのインストール

- ▶ `https://neo4j.com/docs/graph-algorithms/current/graph-algorithms-algo-3.4.7.0.jar` を Neo4j の `plugins` フォルダにおく
- ▶ Neo4j の config file に `dbms.security.procedures.unrestricted=algo.*` を書き込む
- ▶ Neo4j を再起動



Google Page Rank

```
CALL algo.pageRank.stream(null, 'depends',  
    {iterations:20, dampingFactor:0.85})  
YIELD nodeId, score  
MATCH (node) WHERE id(node) = nodeId  
RETURN node.name AS page,score  
ORDER BY score DESC
```

Page rank の結果

page	score
context	4.8682650000000001
hyphen-base	4.6671720000000001
hyph-utf8	4.0754105
kpathsea	1.8529665
plain	0.982524
cm	0.69867500000000002
luatex	0.656277
pdftex	0.621257
collection-basic	0.556997



Betweenness Centrality

```
CALL algo.betweenness.stream(null,'depends',
    {direction:'out'})
YIELD nodeId, centrality
MATCH (pkg) WHERE id(pkg) = nodeId
RETURN pkg.name AS pkg,centrality
ORDER BY centrality DESC;
```

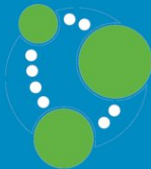
Betweenness の結果

pkg	centrality
collection-basic	1675.4717032967033
collection-latexextra	1212.0
context	947.3333333333334
collection-latex	744.8166666666666
collection-pictures	586.0
collection-mathscience	528.0
collection-plaingeneric	448.5
collection-latexrecommended	412.0
xetex	336.41666666666663



Triangle computation

```
CALL algo.triangleCount.stream(null, 'depends',  
    {concurrency:4})  
YIELD nodeId, triangles, coefficient  
MATCH (p) WHERE id(p) = nodeId  
RETURN p.name AS name, triangles, coefficient  
ORDER BY triangles DESC
```



結論

- ▶ Neo4j で簡単に任意のシステムを表現できる



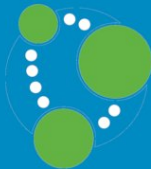
結論

- ▶ Neo4j で簡単に任意のシステムを表現できる
- ▶ 簡単な Cypher で複雑な状況を確認できる



結論

- ▶ Neo4j で簡単に任意のシステムを表現できる
- ▶ 簡単な Cypher で複雑な状況を確認できる
- ▶ 早い（今日だけの準備！）



結論

- ▶ Neo4j で簡単に任意のシステムを表現できる
- ▶ 簡単な Cypher で複雑な状況を確認できる
- ▶ 早い（今日だけの準備！）

興味があれば

<http://www.texlive.info:7474/>

で遊んでいい ;-) (read-only)

結論

- ▶ Neo4j で簡単に任意のシステムを表現できる
- ▶ 簡単な Cypher で複雑な状況を確認できる
- ▶ 早い（今日だけの準備！）

興味があれば

<http://www.texlive.info:7474/>

で遊んでいい ;-) (read-only)

ありがとうございます