

Wprowadzenie do pracy z literaturą naukową: wprowadzenie do LaTeXa

Jan Karwowski, Stanisław Kaźmierczak

October 29, 2021

Outline

L^AT_EX

Szczegółowe użycie

Składanie bibliografii

Narzędzia pomocnicze

Co to jest $\text{\LaTeX}$

$\text{\LaTeX}$ jest zbiorem makr ułatwiających korzystanie z języka $\text{\TeX}$.

Czym w takim razie jest $\text{\TeX}$? $\text{\TeX}$ jest **językiem programowania** stworzonym przez D. Knutha, wyspecjalizowanym do generowania dokumentów do druku.

Prosty dokument w LaTeXu

```
\documentclass[a4paper]{article} % definicja klasy dokumentu
%początek preambuły
\usepackage[utf8]{inputenc}
\usepackage{amssymb}
\usepackage[MeX]{polski}
\title{XYZ}
\author{Jan Karwowski}
%koniec preambuły
\begin{document}
\section{Pierwsza sekcja}
treść sekcji
\end{document}
```

Dokument w LaTeXu możemy podzielić na dwie podstawowe części: preambułę i właściwą treść dokumentu.

Najważniejsze elementy składni

- ▶ Większość tekstu wprowadzanego pomiędzy `\begin{document}` i `\end{document}` będzie po prostu treścią generowanego dokumentu. Uwaga: jest całkiem sporo znaków specjalnych. Pojedyncze znaki nowej linii są ignorowane, podwójne definiują akapity.
- ▶ Makra – w TeXu makra wywoływane są poprzez podanie nazwy makra poprzedzonej backslashem. Jeśli makro przyjmuje jakieś argumenty, to są one podawane w nawiasach klamrowych lub prostokątnych po nazwie makra. Argumenty w nawiasach kwadratowych to argumenty opcjonalne.
- ▶ komentarze: znak `%` oznacza komentarz od bieżącego miejsca do końca linii.
- ▶ otoczenia (environments) – mamy dwa wyspecjalizowane makra `\begin{X}` i `\end{X}`. $\text{\LaTeX}$ pozwala definiować otoczenia, które albo w jakiś sposób zmieniają styl tekstu, albo udostępniają specjalne makra, formatowania, etc.
- ▶ zakresy ograniczone klamerkami: większość wartości zmiennych globalnych w TeXu obowiązuje w obrębie zakresu, np. `{\color{green} abc def}` `ghi` – ustawienie koloru działa dla `abc` i `def`, ale `ghi` jest już poza zakresem.

Klasa dokumentu

Klasa dokumentu definiuje bazowy zestaw makr do używania w danym dokumencie. W większości przypadków będzie to podstawowe formatowanie tekstu i podział na rozdziały/sekcje/ itp. Dostępne elementy różnią się w zależności od klasy. Np. w klasie `book` da się definiować rozdziały, a w klasie `article` tylko sekcje. W klasie `beamer`, która służy do przygotowania prezentacji podstawowym bytem są ramki (`frame`, czyli slajdy), które nie istnieją w innych klasach.

Dokument może mieć dokładnie jedną klasę (`\documentclass` w pierwszej linii). Poza klasami dostarczonymi z daną dystrybucją LaTeXa, często można spotkać klasy przygotowane przez wydawnictwa, przeznaczone do przygotowania dokumentów na daną konferencję/do danego czasopisma.

Podział dokumentu na mniejsze pliki

Są dwa makra pozwalające włączyć treść zewnętrznego pliku:

- ▶ `include` – łamie stronę po zawartości pliku, używa oddzielnych plików pomocniczych dla włączonych zawartości, nie może być zagnieżdżany

- ▶ `input`

W obu przypadkach ścieżki względne we włączonych plikach rozpatrywane są względem głównego pliku $\text{\TeX}$!

Pakiety

Pakiety definiują dodatkowe makra, mamy na przykład pakiety definiujące dodatkowe symbole matematyczne (amssymb), dodatkowe komendy do formatowania tabel (booktabs), kolorowania (xcolor), itd.

Do załączania pakietów służy makro `\usepackage[]{}{}`. Obowiązkowy argument (w nawiasach klamrowych) to nazwa pakietu, a opcjonalny to opcje pakietu.

W momencie kiedy zaczynamy korzystać z pakietów, czyli praktycznie zawsze, należy pamiętać, że pod względem technologicznym $\text{\LaTeX}$ jest bardzo ubogi. wszelkiego rodzaju konfiguracja odbywa się z użyciem ustawiania bezargumentowych makr (odpowiednik zmiennych globalnych), co oznacza, że pakiety mogą skutecznie przeszkadzać sobie nawzajem, szczególnie jeśli zakres ich odpowiedzialności pokrywa się ze sobą. W przypadku dziwnych efektów zawsze warto poszukać informacji o niepożądanych interakcjach między używanymi pakietami.

Repozytorium pakietów

ctan.org: The Comprehensive T_EX Archive Network
Zawiera też dokumentację pakietów

Struktura dokumentu

Istotną cechą LaTeXa jest to, że promuje on definiowanie struktury dokumentu, a nie wyglądu.

W większości klas mamy dostępne następujące makra do definiowania kolejnych poziomów nagłówków:

- ▶ `\section{Tytuł sekcji}`
- ▶ `\subsection{Tytuł sekcji}`
- ▶ `\subsubsection{Tytuł sekcji}`
- ▶ `\paragraph{Tytuł akapitu}`

W większości przypadków sekcje są numerowane, np, trzecia podsekcja w drugiej sekcji ma numer 2.3. Akapity zwykle nie są numerowane. Style numeracji można oczywiście konfigurować.

Za pomocą makra `\tableofcontents` można wstawić spis treści.

Odnosińki

Powodem dla którego w dokumentach użytkowych numeruje się poszczególne elementy (np. sekcje) jest łatwość odnoszenia się to tych bytów poprzez wskazanie numerów. Ręczne wpisywanie numerów – problem przy zmianie struktury (co powoduje zmianę numeracji).

L^AT_EX ma makra:

- ▶ `\label{nazwa}` – do zaznaczania punktu, do którego będziemy się odnosić, odpowiednik `<a name="nazwa">` w htmlu.
- ▶ `\ref{nazwa}` – generuje odośnik – numer bytu, w którym jest `\label{nazwa}`. Jest to numer ostatniego numerowanego bytu widocznego w danym zakresie.
- ▶ `\pageref{nazwa}` – numer strony, gdzie jest `\label{nazwa}`.

Przykład dokumentu z etykietami

```
% profanacja wieszczą
\section{Wpłynąłem na}\label{sec:wpl}% 1
suchego przestwór oceanu, Wóz nurza się w zieloność
\subsection{i jak łódka brodzi,}
Śród fali łąk
\label{sec:ija} % 1.1
szumiących,
\label{sec:ija2} % 1.1
\subsection{śród kwiatów powodzi,}
\subsubsection{Omijam}
    koralowe ostrowy burzanu.\label{sec:omi}% 1.2.1
\section{Już mrok zapada,}
    nigdzie drogi ni kurhanu;\label{sec:juz} % 2
```

Konwencje nadawania etykiet

- ▶ Etykiety są po prostu napisami, dowolnymi.
- ▶ Przyjęło się, żeby łatwiej nawigować po dokumencie stosować prefiksy
 - ▶ etykiety sekcji zwykle zaczyna się od `sec:`, np. `sec:introduction`
 - ▶ etykiety tabel `tab:`
 - ▶ etykiety obrazków (figure) `fig:`
 - ▶ etykiety równań `eq:`
 - ▶ itd.
- ▶ Tych nazw nie widać w łatwy sposób w wyjściu w formie pdf

Proces kompilacji

- ▶ Kompilacja TeXa jest jednoprzebiegowa. To znaczy, że czytany jest kod, a potem na bieżąco generowane jest wyjście. W przypadku odnośników do elementów, które znajdują się dalej w tekście (do przodu), pojawia się problem.
- ▶ Uruchomienie TeXa poza właściwym plikiem wynikowym generuje pliki .aux i .toc (i w zależności od użytych pakietów może generować jeszcze inne). Pliki pomocnicze zawierają informacje, które mogą się przydać przy kolejnym przebiegu kompilacji, np. pozycje etykiet w poprzednim przebiegu. Te informacje są wykorzystywane przez makro `\ref`.
 - ▶ dwukrotny przebieg może nie wystarczyć, bo wstawienie odnośnika może spowodować przeniesienie jakiegoś kawałka tekstu na kolejną stronę.
 - ▶ wyjście przebiegu TeXa zawiera informację, o tym, czy od ostatniego przebiegu zmieniły się pozycje etykiet.

Wyjście kompilacji

- ▶ Tradycyjnym formatem wyjściowym TeXa jest format DVI (device independent format).
- ▶ Większość dystrybucji TeXa pozwala również generować PostScript i PDF.
- ▶ Typowe dystrybucje mają trzy różne polecenia: `latex`, `pslatex`, `pdflatex`, które generują wyjście w różnych formatach.

Dystrybucje TeXa

Dygresja.

Czym są dystrybucje TeXa. Nie ma jednej centralnej wersji kompilatora TeXa lub LaTeXa i zbioru różnego rodzaju pakietów. Są różne zestawy (potencjalnie z różnymi kompilatorami), które zawierają różne zestawy i wersje pakietów.

Najpopularniejsze:

- ▶ MikTeX
- ▶ TeXlive

Matematyka

L^AT_EX jest szczególnie znany z tego, że da się w nim składać złożone wzory matematyczne w sposób ogarnialny dla człowieka.

Trzy podstawowe sposoby wprowadzania matematyki:

wzór wewnątrz tekstu jest po prostu traktowany jak pozostałe słowa a tekście, jest częścią linii i może być łamany automatycznie na wiele linii (inline math). Może być wprowadzony na dwa sposoby, albo pomiędzy dolarami: $a+b$ albo pomiędzy makrami $($ i $)$: $(a+b)$.

wzór wyróżniony (display math). Jest w oddzielnej przestrzeni między liniami. Silnik składający ten wzór działa też trochę inaczej, m.in indeksy dolne i górne są bardziej rozsunięte, większe są odstępy przy kresce ułamkowej, w przypadku sum i podobnych granice są pod symbolem. Wprowadzany albo z użyciem $\$$ albo z użyciem makr $[$ i $]$, ewentualnie otoczenie `equation*`.

wzór numerowany szczególny przypadek poprzedniego, wzór dodatkowo ma numer (i można mu zdefiniować etykietę). Otoczenie `equation`.

amsmath jest popularnym pakietem rozszerzającym podstawowe możliwości składania wzorów matematycznych, właściwie każda praca, w której są wzory, używa tego pakietu.

Pozwala między innymi na:

- ▶ eleganckie wyrównywanie układów równań
- ▶ łamanie równań na kilka linii
- ▶ możliwość definiowania własnych nazw operatorów w stylu min i max
- ▶ ...

Dodatkowo: amssymb – rozszerzony zbiór symboli

Otoczenie verbatim

Czasami wygodnie jest wprowadzić fragment tekstu, bez interpretacji znaków specjalny. Tekst w otoczeniu verbatim jest składany fontem stałej szerokości, a znaki specjalne nie są interpretowane (podobnie jak `ıpreł` w `htmlu`)

Elementy pływające

Poza tekstem, często będziecie umieszczać rysunki, tabele lub podobne elementy.

- ▶ te elementy mają kształt prostokąta
- ▶ nie mieszczą się w linii
- ▶ umieszczenie ich w środku zdania spowodowałoby zaburzenie płynności czytania tekstu
- ▶ $\text{\LaTeX}$ potrafi je umieścić w pobliżu miejsca, gdzie element został umieszczony w kodzie (czasami na innej stronie)
- ▶ Makro `\clearpage` powoduje, że wszystkie nieumieszczone elementy pływające zostaną umieszczone przed następującym dalej tekstem.
- ▶ pakiet `float` umożliwia dokładniejszą konfigurację
- ▶ są pakiety do robienia podrysunków

Rysunki

Otoczenie float

```
\begin{figure}[t]  
% kod realizujący rysunek  
\caption{Schemat połączeń w sieci wewnętrznej}  
\label{fig:conn-scheme}  
\end{figure} %tu kończy się zakres obowiązywania caption
```

Label zawsze po caption

Argument opcjonalny, tu t, definiuje preferowane umieszczenie elementu, tu top – na górze strony.

Rysunki – wstawianie

Do wstawiania rysunków z zewnętrznych plików zwykle używa się pakietu `graphicx`

```
\usepackage{graphicx}  
%...
```

```
\includegraphics[width=\columnwidth]{rysunek}
```

- ▶ Zwykle `includegraphics` umieszcza się wewnątrz `figure`.
- ▶ Rozszerzenia pliku
 - ▶ Podając ścieżkę bez rozszerzenia wybiera się format preferowany dla danego formatu wyjściowego (nie każde wyjście obsługuje każdy format wejściowy). Można mieć kilka nazw z tym samym rozszerzeniem i wybrać preferowane.
 - ▶ Można podać ścieżkę z rozszerzeniem.
 - ▶ Dla wyjścia PDF najlepiej mieć grafiki wektorowe w PDF
- ▶ Skalowanie nie jest jedyną możliwą wartością opcjonalnego argumentu.

Tabele

Pływające otoczenie do tabel `table`. Oddzielna numeracja od rysunków. Tabele zwykle mają podpis (`caption`) nad tabelą, a nie pod.

Składanie tabeli: otoczenie `tabular`.

```
\begin{tabular}{cr|rl|c}  
a&b&c&d&e\\  
\hline  
f&g&h&i&j\\  
a&b&c&d&e\\  
\end{tabular}
```

Argument otoczenia specyfikuje kolumny, litera oznacza wyrównanie (do lewej, do prawej, do środka, są też bardziej zaawansowane specyfikacje, np. odstępów między kolumnami). Pionowa kreska oznacza kreskę między kolumnami.

W środowiskach, które preferują użycie pakietu `booktabs`, wstawianie pionowych kresek jest bardzo nieeleganckie.

Pseudokody

Jest dostępnych wiele pakietów do formatowania pseudokodów (niekompatybilnych ze sobą nawzajem), na przykład:

- ▶ `algorithmic`
- ▶ `algorithm2e`
- ▶ `algorithmicx`

Ważniejsze funkcjonalności:

- ▶ formatowanie funkcji, słów kluczowych, itd.
- ▶ różne formy wyróżniania bloków
- ▶ numerowanie linii (z etykietami i możliwością odnośników)

Kody programów

Pakiet `listings` pozwala wstawić kod programu, obsługuje też kolorowanie składni.

Hyperref

Pakiet wpływający na wyjście w formacie PDF, dodający odnośniki i metadane.

Nagłówki i stopki

- ▶ Zwykle domyślny format nagłówków i stopek jest definiowany przez klasę dokumentu.
- ▶ W prostych klasach będą numery stron, w bardziej zaawansowanych mogą się pojawiać takie elementy, jak tytuł dokumentu, tytuł rozdziału, itp.
- ▶ Pakiet `fancyhdr` pozwala na łatwą konfigurację nagłówków i stopek.

Grafika wektorowa w ramach dokumentu

Poza włączaniem grafik z plików zewnętrznych, są pakiety, które pozwalają rysować w użyciem makr $\text{T}_{\text{E}}\text{X}$. Popularne to:

- ▶ pstricks
- ▶ tikz

Komunikaty kompilacji

Warto czytać ostrzeżenia kompilacji (oczywiście)!

- ▶ Odniesienia do brakujących etykiet (w wyjściu zastąpione znakami zapytania)
- ▶ Brakujące wpisy w bibliografii (omówione dalej)
- ▶ Problemy z łamaniem linii:
 - ▶ Overfull – $\text{T}_{\text{E}}\text{X}$ nie znalazł dobrego sposobu na złamanie linii i fragment jest za marginesem. Może dotyczyć też zbyt szerokich tabel i obrazków.
 - ▶ Underfull – punkt złamania linii był na tyle wcześnie, że tekst jest nienaturalnie rozstrzelony.
- ▶ Informacja czy zmieniły się pozycje etykiet i czy potrzebny jest kolejny przebieg.

Pakiety do zarządzania bibliografią

Pakiety, które służą do generowania listy cytowań na końcu dokumentu wraz z odnośnikami:

- ▶ bibtex (najstarszy i najpopularniejszy)
 - ▶ Tak naprawdę dwie części
 - ▶ Pakiet do LaTeXa
 - ▶ Narzędzie do przetwarzania bazy bibliografi w kod LaTeXa
- ▶ biblatex
- ▶ natbib

Dwa niżej potrafią używać baz generowanych przez konsolowe narzędzie bibtex, jak i baz generowanych przez biber (inne narzędzie, nowsze).

Struktura pliku bibtex

Wpisy postaci:

```
@InProceedings{KarwowskiMZGA19,  
  author      = {Jan Karwowski and Jacek Ma{\`n}dziuk and Adam {\.Z}ychow  
and Filip Grajek and Bo An},  
  title       = {A Memetic Approach for Sequential Security Games on  
a Plane with Moving Targets},  
  year        = 2019,  
  booktitle   = {The Thirty-Third {AAAI} Conference on Artificial  
Intelligence, {AAAI} 2019, The Thirty-First  
Innovative Applications of Artificial Intelligence  
Conference, {IAAI} 2019, The Ninth {AAAI} Symposium  
on Educational Advances in Artificial Intelligence,  
{EAAI} 2019, Honolulu, Hawaii, USA, January 27 -  
February 1, 2019},  
  pages       = {970-977},  
  doi         = {10.1609/aaai.v33i01.3301970},  
  url         = {https://doi.org/10.1609/aaai.v33i01.3301970}
```

Wykorzystanie BibTeXa

W przykładowym dokumencie

Sprawdzanie częstych błędów

- ▶ lacheck – czasami zdarza się zapomnieć zamknąć jakąś klamerkę, środowisko, itp. Interpreter LaTeXa jest mało domyślny i zwykle rzuca błąd na końcu dokumentu. lacheck ma heurystyki, które nieźle wskazują miejsce problemu.
- ▶ chktex – zestaw wielu testów, głównie związany z typografią i wyglądem.

Automatyzacja powtórzeń kompilacji

Na podstawie logów i dat modyfikacji można automatycznie stwierdzić czy konieczna jest ponowna kompilacja lub przebudowanie BibTeXa. Są narzędzia, które automatyzują cały przebieg, na przykład rubber.