

LUATEX

ANNO 2008

A SHORT HISTORY

- * THE TEAM: TACO HOEKWATER, HARTMUT HENKEL, HANS HAGEN
- * THE STARTING POINT: A STRIPPED DOWN $\text{PDF}\TeX$
- * THE GOODIES: THE ALREADY PRESENT $\epsilon\TeX$
- * THE NEW KID: A Lua SUBSYSTEM
- * THE CATALYST: IDRIS' ORIENTAL \TeX PROJECT
- * THE ODD KID: SOME ALEPH FUNCTIONALITY NEEDED BY IDDS
- * THE CHALLENGE: REWRITE MANY PARTS IN C AND PROVIDE APIs
- * THE FUTURE: CONVERT THE PASCAL SOURCE INTO CWEB AND GLUE WITH Lua
- * THE CURRENT STATE: THE SECOND BETA TEST CYCLE (THIRD BETA BY THE END OF 2008)
- * THE NEXT PAGES: ONLY A QUICK OVERVIEW (NO TIME FOR DETAILS)

WIDE AND VIRTUAL FONTS

- ★ TFM LOADING HAS BEEN GENERALIZED (VIA LUA TABLE)
- ★ OPENTYPE AND TYPE1 LOADERS HAVE BEEN ADDED
- ★ LOADING FONTS CAN BE BROUGHT UNDER LUA CONTROL
- ★ THERE IS NO LONGER GLOBAL SHADING OF FONT DATA
- ★ AS SUCH, HZ AND PROTRUDING VECTORS ARE NOW PER FONT
- ★ EXPERIMENTAL PDP_{TEX} FEATURES HAVE BEEN REMOVED
- ★ GLYPH NODES CAN BE ACCESSED AND CAN CARRY MORE INFORMATION
- ★ THERE IS PDF OUTPUT FOR WIDE FONTS (2³¹ CHARACTERS)
- ★ VIRTUAL FONTS CAN BE BUILT (IN A STILL SOMEWHAT EVOLVING INTERFACE)

TOKENIZING AND NODE LISTS

- * THERE ARE CALLBACKS FOR VARIOUS INTERNAL FUNCTIONS
- * MOST (AND SOON ALL) MEMORY MANAGEMENT HAS BEEN REWRITTEN
- * INPUT CAN BE INTERCEPTED WHENEVER T_EX WANTS TO READ SOMETHING
- * TOKENIZATION CAN BE INTERCEPTED AND EXTENDED (SLOW)
- * AT VARIOUS STAGES THE NODE LISTS CAN BE INTERCEPTED (FAST)
- * NODES CAN CARRY MANY ATTRIBUTES (TO BE SET FROM T_EX)
- * CALLBACKS DEALING WITH NODE LISTS CAN USE INFORMATION STORED IN ATTRIBUTES

PARAGRAPH AND PAGE BUILDING

- ★ THE VARIOUS STAGES OF PARAGRAPH BUILDING ARE SEPARATED
- ★ HYPHENATION, KERNING AND LIGATURE BUILDING CAN BE OVERLOADED
- ★ HYPHENATION PATTERNS AND EXCEPTIONS CAN BE LOADED AT RUNTIME
- ★ THERE ARE A FEW HELPERS FOR MANIPULATION OF NODE LISTS
- ★ THE MAIN VERTICAL LIST BUILDING CAN BE INFLUENCED (WILL BE EXTENDED)
- ★ NEXT ON THE AGENDA IS OPENING UP MATH AND ALIGNMENTS (END 2008)
- ★ LATER IN THE PROJECT THE OUTPUT ROUTINE RELATED MECHANISMS WILL BE OPENED UP (2009)

GRAPHIC CAPABILITIES AND BACKEND

- ★ THERE IS A LIBRARY FOR IMAGE IDENTIFICATION AND INCLUSION
- ★ AT SOME POINT OBJECT MANAGEMENT, ANNOTATION SUPPORT AND OTHER BACKEND ISSUES WILL BE UNDER LUA CONTROL
- ★ THE ALPHA VERSION OF THE METAPOST LIBRARY IS INCLUDED (MULTIPLE INSTANCES ARE POSSIBLE)
- ★ THE VIRTUAL FONTS 'COMMANDS' ARE BEING PREPARED FOR RUNTIME FONT GENERATION
- ★ LATER IN THE PROJECT WE WILL BE ABLE TO CONSTRUCT FONTS AT RUNTIME (BY MANIPULATING SO CALLED CHARSTRINGS)

TEX PRIMITIVES AND INTERNALS

- ★ WE HAVE (NAMED) ACCESS TO AND CONTROL OVER ALL REGISTERS
- ★ THERE WILL BE (MORE THAN CURRENTLY) ACCESS TO SPECIAL REGISTERS
- ★ SOME CURRENTLY INACCESSIBLE INTERNAL QUANTITIES THAT CONTROL TYPESETTING WILL BE ACCESSIBLE
- ★ AT SOME POINT WE WILL REINVENT BOX HANDLING (E.G. BETTER CONTROL OVER LOCAL AND GLOBAL BEHAVIOUR FROM LUA)

CONTEXT MkIV

ANNO 2008

A SHORT HISTORY

- ★ MKIV IS A BRANCH OF CONTEXT TUNED FOR L^AT_EX
- ★ IT IS MOSTLY BACKWARD COMPATIBLE BUT PROVIDES MORE
- ★ DURING DEVELOPMENT WE USE IT IN CRITICAL PRODUCTION ENVIRONMENTS
- ★ ALL SUPPORT SCRIPTS ARE DONE IN L^UA, SO NO MORE DEPENDENCIES
- ★ THERE ARE TWO NEW CORE UTILITIES: LUATOOLS AND MTXDUN
- ★ WE DO LOTS OF TESTING AND BENCHMARKING AND KEEP TRACK OF HISTORY

THE IMPLICATIONS

- ★ READING FROM FILES AND LOGGING IS NOW UNDER LUA CONTROL
- ★ INPUT ENCODING (REGIMES) HAVE BEEN SIMPLIFIED AND FONT ENCODING IS GONE
- ★ CHARACTER AND MATH DEFINITIONS ARE SIMPLIFIED AND MORE EFFICIENT
- ★ VERBATIM AND BUFFERING IS NOW DONE PURELY IN LUA AS WILL BE PRETTY PRINTING
- ★ HYPHENATION PATTERNS ARE LOADED AT DUNTIME AND WE PROVIDE GOODIES LIKE VISUALIZING SPELLING ERRORS
- ★ GRAPHIC FILE IDENTIFICATION, CONVERSION AND INCLUSION IS DONE IN LUA
- ★ MULTIPASS DATA IS KEPT IN LUA TABLES
- ★ INDEX SORTING HAPPENS AT RUNTIME INSTEAD OF BY AN EXTERNAL SCRIPT
- ★ METAPOST CONVERSION HAS BEEN REDONE AND EMBEDDED METAPOST IS NOW INSTANTANEOUS

MODE IMPLICATIONS

- ★ WE'RE EXPERIMENTING WITH A NEW AND MORE FLEXIBLE XML PROCESSING MODEL (OF COURSE IN PURE LUA)
- ★ MATHML SUPPORT IS BEING ADAPTED USING THIS NEW MODEL WHICH GIVES A SPEED-UP
- ★ SECTIONING AND LIST PROCESSING WILL BE REIMPLEMENTED AND MADE MORE FLEXIBLE (HANDY FOR XML PROCESSING)
- ★ CROSS REFERENCING (ESPECIALLY THE COMPLEX PART THAT INVOLVES PARSING) WILL BE REIMPLEMENTED
- ★ CONTEXT WILL BE MORE MODULARIZED SO THAT USERS CAN MAKE SPECIAL VERSIONS FOR AUTOMATED WORKFLOWS
- ★ WE WILL EXPERIMENT WITH MORE EFFICIENT PROCESSING (BASICALLY ONE DUN)

FONTS

- ★ ALL FONT HANDLING IS DONE IN L^UA, INCLUDING O^PENT^ETYPE FEATURES
- ★ WE AIM AT MORE ROBUST CJK SUPPORT
- ★ WE WILL PROVIDE MORE CONTROL OVER TYPESETTING AND FONTS AS DEMANDED BY O^RIENTAL T^EX USAGE
- ★ USERS CAN CONSTRUCT VIRTUAL FONTS AND WE PROVIDE ADDITIONAL FEATURES BASED IN THIS TECHNOLOGY
- ★ CJK SUPPORT IS UNDER RECONSTRUCTION AND WILL BE MORE FLEXIBLE AND CONFIGURABLE
- ★ ADVANCED FEATURES LIKE FONT EXPANSION (HZ) AND PROTRUDING HAVE BEEN REIMPLEMENTED AND ARE AVAILABLE AS FEATURES

ATTRIBUTES

- ★ MOST SCOPE (GROUPING) RELATED FEATURES WILL BE REIMPLEMENTED USING ATTRIBUTES
- ★ AS AN EXPERIMENT, COLOR SUPPORT AND SPECIAL RENDERING OPTIONS (E.G. PDF SPECIFIC) ALREADY HAVE BEEN REWRITTEN
- ★ THIS DEMANDS EXTENSIVE POSTPROCESSING DURING SHIFOUT WHICH IS ALSO DONE IN LUA
- ★ THE REPERTOIRE OF MANIPULATION FEATURE IN CONTEXT HAS BEEN EXTENDED (THINK OF MORE ROBUST CHARACTER CASE MANAGEMENT, HYPHENATING URL'S, ETC.)
- ★ WE HAVE (EXPERIMENTAL) MORE ADVANCED VERTICAL SPACE MODELS (GLUE/PENALTY)
- ★ IN DUE TIME WE WILL PROVIDE A PLUGIN MODEL FOR USERS WHO WANT TO MANIPULATE THE INPUT, FONTS AND NODE LISTS USING ATTRIBUTES

THE RUNNER: MTXRUN

- ★ THIS SCRIPT PERMITS OTHER SCRIPT TO RUN WITHIN MkIV FILE AND FEATURE SPACE (THERE IS NO NEED TO LOAD EXTRA LIBRARIES)
- ★ IN SOME WAY IT IS A REPLACEMENT FOR TEXMPSTART IN MkII AND AS SUCH IT ALSO PROVIDES THE FEW EXTRA BUILT IN FEATURES (DOCUMENT VIEWER, EDIT TREE FILES, ETC.)
- ★ AN EXAMPLE OF A SCRIPT IS 'CONTEXT' WHICH MANAGING A \TeX RUN (MULTIPLE RUNS, EXTRA ACTIONS), AND IS A REPLACEMENT FOR \TeXEXEC
- ★ ANOTHER SCRIPT MAINTAINS THE FONT DATABASE AND PROVIDES USERS A WAY TO IDENTIFY FONT NAMES
- ★ OTHER SCRIPTS DEAL WITH CONVERSION FORM OLD FASHIONED \TeX INPUT, CACHE MANAGEMENT, UPDATING CONTEXT MINIMALS, WEB BASED PUBLISHING