Machine Learning and Configurable Systems: A Gentle Introduction

(tutorial at SPLC'19)

Hugo Martin, Juliana Alves Pereira,

Paul Temple, Mathieu Acher

https://github.com/VaryVary/

How to ensure that all Linux kernel configurations build?

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

KConfig file

```
[...]

config PRINTK
	default y
	bool "Enable support for printk" if EXPERT
	select IRQ_WORK
	help
	  This option enables normal printk support. Removing it
	  eliminates most of the message strings from the kernel image
	  and makes the kernel more or less silent. As this makes it
	  very difficult to diagnose system problems, saying N here is
	  strongly discouraged.

config PRINTK_NMI
	def_bool y
	depends on PRINTK
	depends on HAVE_NMI

config BUG
	bool "BUG() support" if EXPERT
	default y
	help
	  Disabling this option eliminates support for BUG and WARN, reducing
	  the size of your kernel image and potentially quietly ignoring
	  numerous fatal conditions. You should only consider disabling this
	  option for embedded systems with no facilities for reporting errors.
	  Just say Y.

config ELF_CORE
	depends on COREDUMP
	default y
	bool "Enable ELF core dumps" if EXPERT
	help
	  Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
	bool "Enable AIO support" if EXPERT
	default y
	help
	  This option enables POSIX asynchronous I/O which may by used
	  by some high performance threaded applications. Disabling
	  this option saves about 7k.

[...]
```

Given a configuration, what's the size of a Linux kernel?

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

KConfig file

```
[...]

config PRINTK
        default y
        bool "Enable support for printk" if EXPERT
        select IRQ_WORK
        help
          This option enables normal printk support. Removing it
          eliminates most of the message strings from the kernel image
          and makes the kernel more or less silent. As this makes it
          very difficult to diagnose system problems, saying N here is
          strongly discouraged.

config PRINTK_NMI
        def_bool y
        depends on PRINTK
        depends on HAVE_NMI

config BUG
        bool "BUG() support" if EXPERT
        default y
        help
          Disabling this option eliminates support for BUG and WARN, reducing
          the size of your kernel image and potentially quietly ignoring
          numerous fatal conditions. You should only consider disabling this
          option for embedded systems with no facilities for reporting errors.
          Just say Y.

config ELF_CORE
        depends on COREDUMP
        default y
        bool "Enable ELF core dumps" if EXPERT
        help
          Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
        bool "Enable AIO support" if EXPERT
        default y
        help
          This option enables POSIX asynchronous I/O which may by used
          by some high performance threaded applications. Disabling
          this option saves about 7k.

[...]
```

How to ensure that all Linux kernel configurations boot?

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

KConfig file

```
[...]

config PRINTK
	default y
	bool "Enable support for printk" if EXPERT
	select IRQ_WORK
	help
	  This option enables normal printk support. Removing it
	  eliminates most of the message strings from the kernel image
	  and makes the kernel more or less silent. As this makes it
	  very difficult to diagnose system problems, saying N here is
	  strongly discouraged.

config PRINTK_NMI
	def_bool y
	depends on PRINTK
	depends on HAVE_NMI

config BUG
	bool "BUG() support" if EXPERT
	default y
	help
	  Disabling this option eliminates support for BUG and WARN, reducing
	  the size of your kernel image and potentially quietly ignoring
	  numerous fatal conditions. You should only consider disabling this
	  option for embedded systems with no facilities for reporting errors.
	  Just say Y.

config ELF_CORE
	depends on COREDUMP
	default y
	bool "Enable ELF core dumps" if EXPERT
	help
	  Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
	bool "Enable AIO support" if EXPERT
	default y
	help
	  This option enables POSIX asynchronous I/O which may by used
	  by some high performance threaded applications. Disabling
	  this option saves about 7k.

[...]
```

Given a configuration, what's the boot time of a Linux kernel?

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

KConfig file

```
[...]

config PRINTK
        default y
        bool "Enable support for printk" if EXPERT
        select IRQ_WORK
        help
          This option enables normal printk support. Removing it
          eliminates most of the message strings from the kernel image
          and makes the kernel more or less silent. As this makes it
          very difficult to diagnose system problems, saying N here is
          strongly discouraged.

config PRINTK_NMI
        def_bool y
        depends on PRINTK
        depends on HAVE_NMI

config BUG
        bool "BUG() support" if EXPERT
        default y
        help
          Disabling this option eliminates support for BUG and WARN, reducing
          the size of your kernel image and potentially quietly ignoring
          numerous fatal conditions. You should only consider disabling this
          option for embedded systems with no facilities for reporting errors.
          Just say Y.

config ELF_CORE
        depends on COREDUMP
        default y
        bool "Enable ELF core dumps" if EXPERT
        help
          Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
        bool "Enable AIO support" if EXPERT
        default y
        help
          This option enables POSIX asynchronous I/O which may by used
          by some high performance threaded applications. Disabling
          this option saves about 7k.

[...]
```

Configurator

```
.config - Linux/x86 4.2.0 Kernel Configuration
                Linux/x86 4.2.0 Kernel Configuration
  Arrow keys navigate the menu.  <Enter> selects submenus ---> (or empty submenus
  ----).  Highlighted letters are hotkeys.  Pressing <Y> includes, <N> excludes, <M>
  modularizes features.  Press <Esc><Esc> to exit, <?> for Help, </> for Search.
  Legend: [*] built-in  [ ] excluded  <M> module  < > module capable

      [*] 64-bit kernel
          General setup  --->
      [*] Enable loadable module support  --->
      [*] Enable the block layer  --->
          Processor type and features  --->
          Power management and ACPI options  --->
          Bus options (PCI etc.)  --->
          Executable file formats / Emulations  --->
      [*] Networking support  --->
          Device Drivers  --->
          Firmware Drivers  --->
          File systems  --->
          Kernel hacking  --->
          Security options  --->
      -*- Cryptographic API  --->
      [*] Virtualization  --->
          Library routines  --->

          <Select>    < Exit >    < Help >    < Save >    < Load >
```

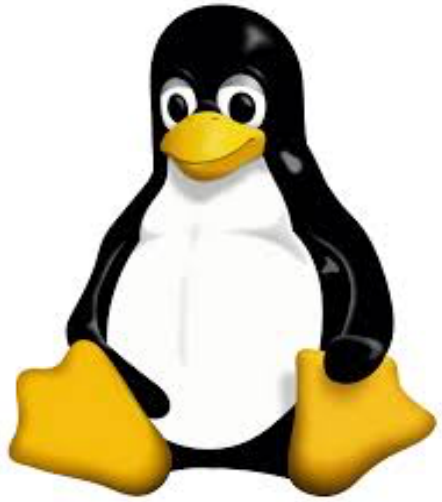
You cannot exhaustively measure all configurations

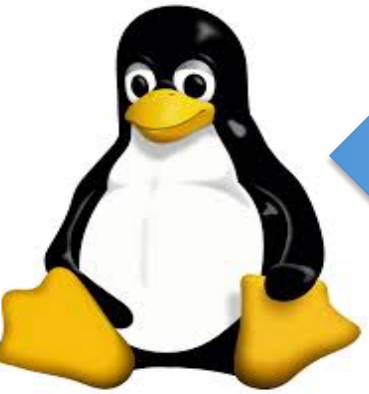
Learning out of a sample

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

KConfig file

```
[...]

config PRINTK
        default y
        bool "Enable support for printk" if EXPERT
        select IRQ_WORK
        help
          This option enables normal printk support. Removing it
          eliminates most of the message strings from the kernel image
          and makes the kernel more or less silent. As this makes it
          very difficult to diagnose system problems, saying N here is
          strongly discouraged.

config PRINTK_NMI
        def_bool y
        depends on PRINTK
        depends on HAVE_NMI

config BUG
        bool "BUG() support" if EXPERT
        default y
        help
          Disabling this option eliminates support for BUG and WARN, reducing
          the size of your kernel image and potentially quietly ignoring
          numerous fatal conditions. You should only consider disabling this
          option for embedded systems with no facilities for reporting errors.
          Just say Y.

config ELF_CORE
        depends on COREDUMP
        default y
        bool "Enable ELF core dumps" if EXPERT
        help
          Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
        bool "Enable AIO support" if EXPERT
        default y
        help
          This option enables POSIX asynchronous I/O which may by used
          by some high performance threaded applications. Disabling
          this option saves about 7k.

[...]
```

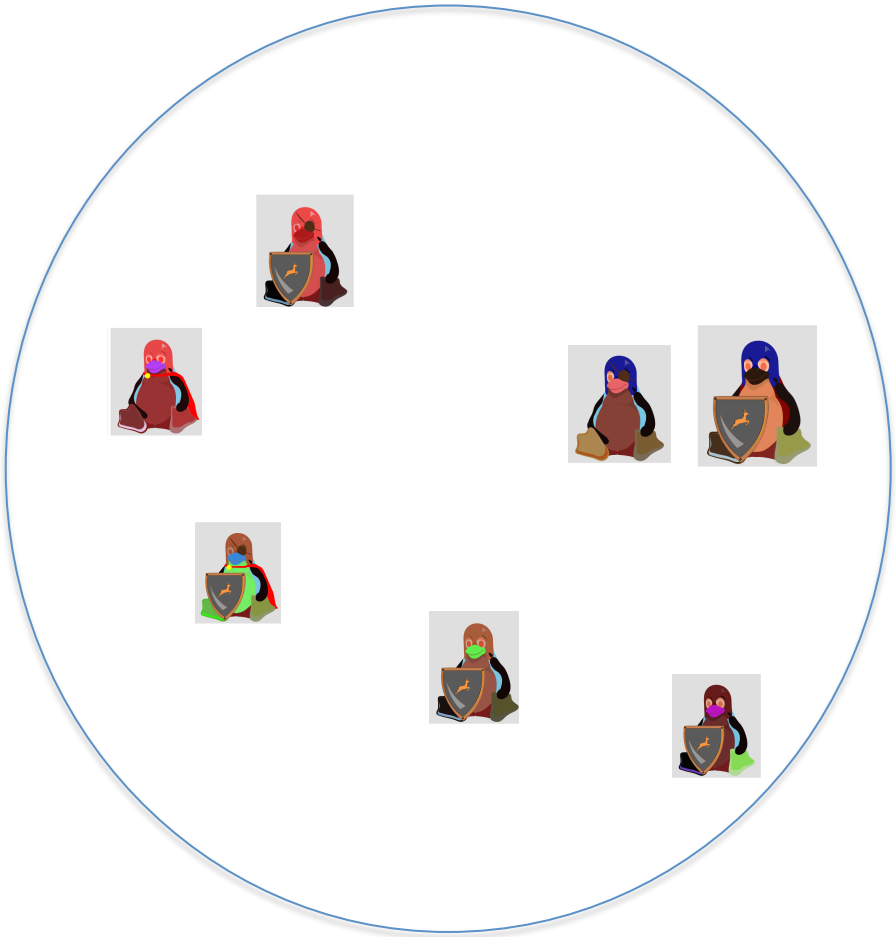




Insights

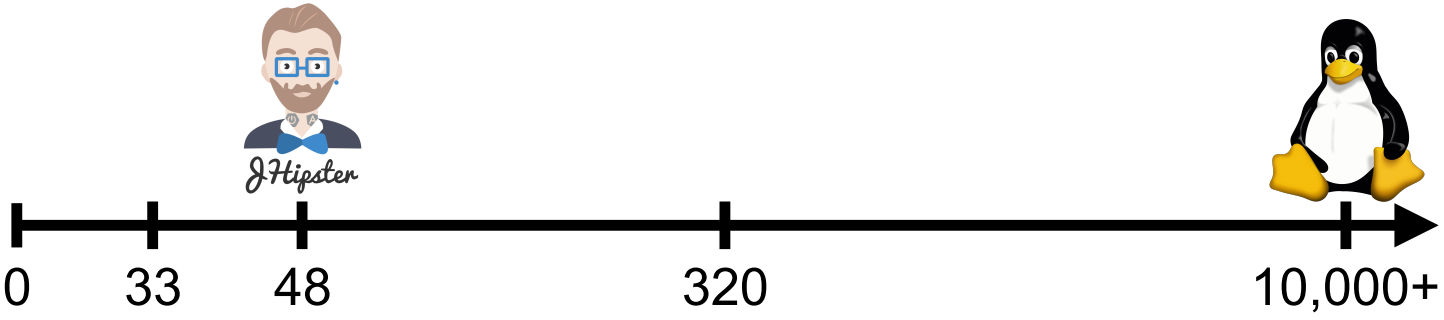
Prediction

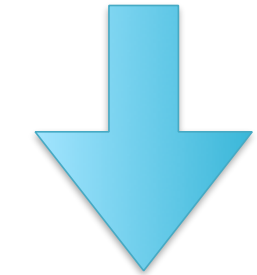
Training set (sample)



Unmeasured variants (testing sample)

A Universe of Options

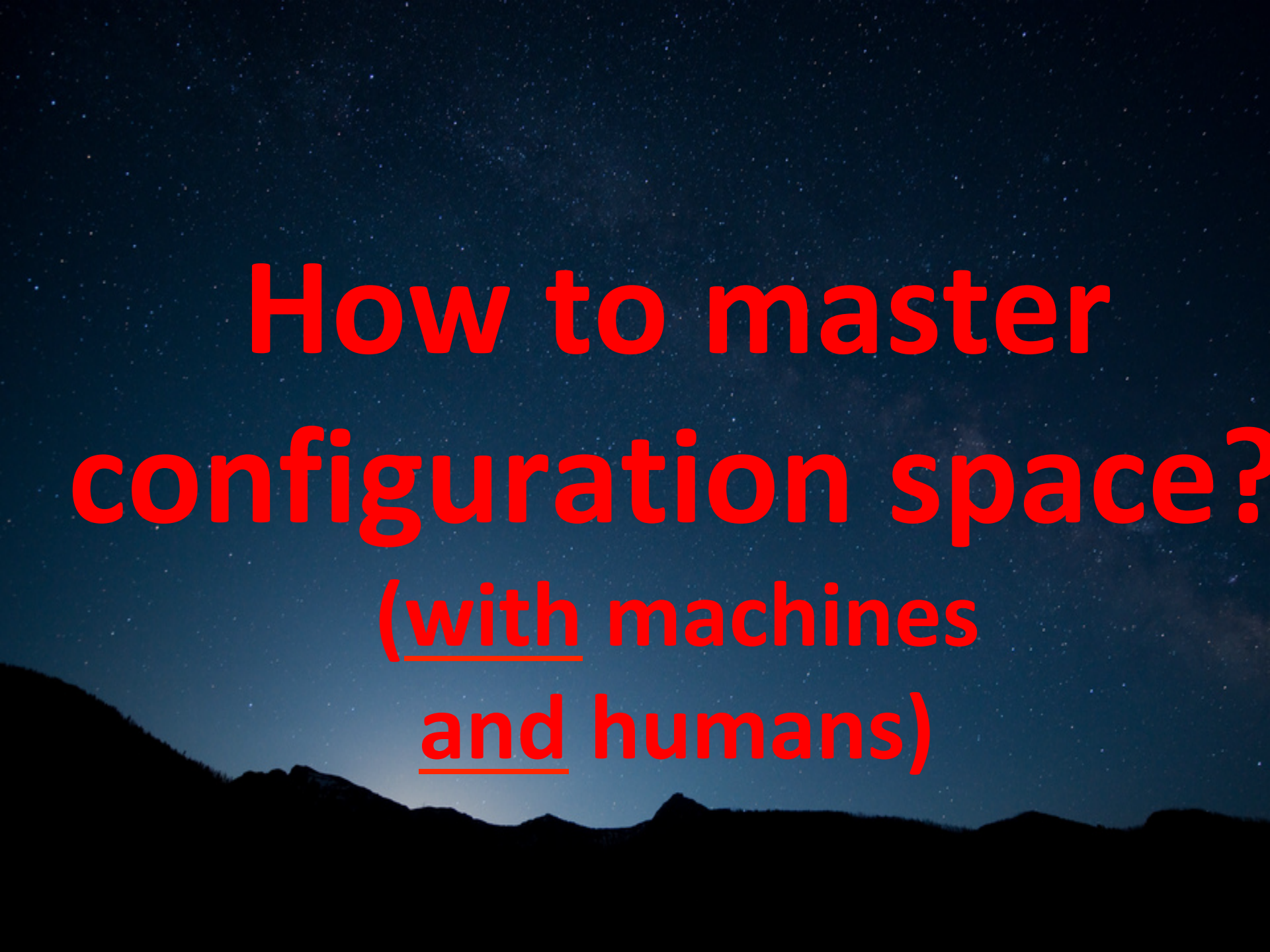


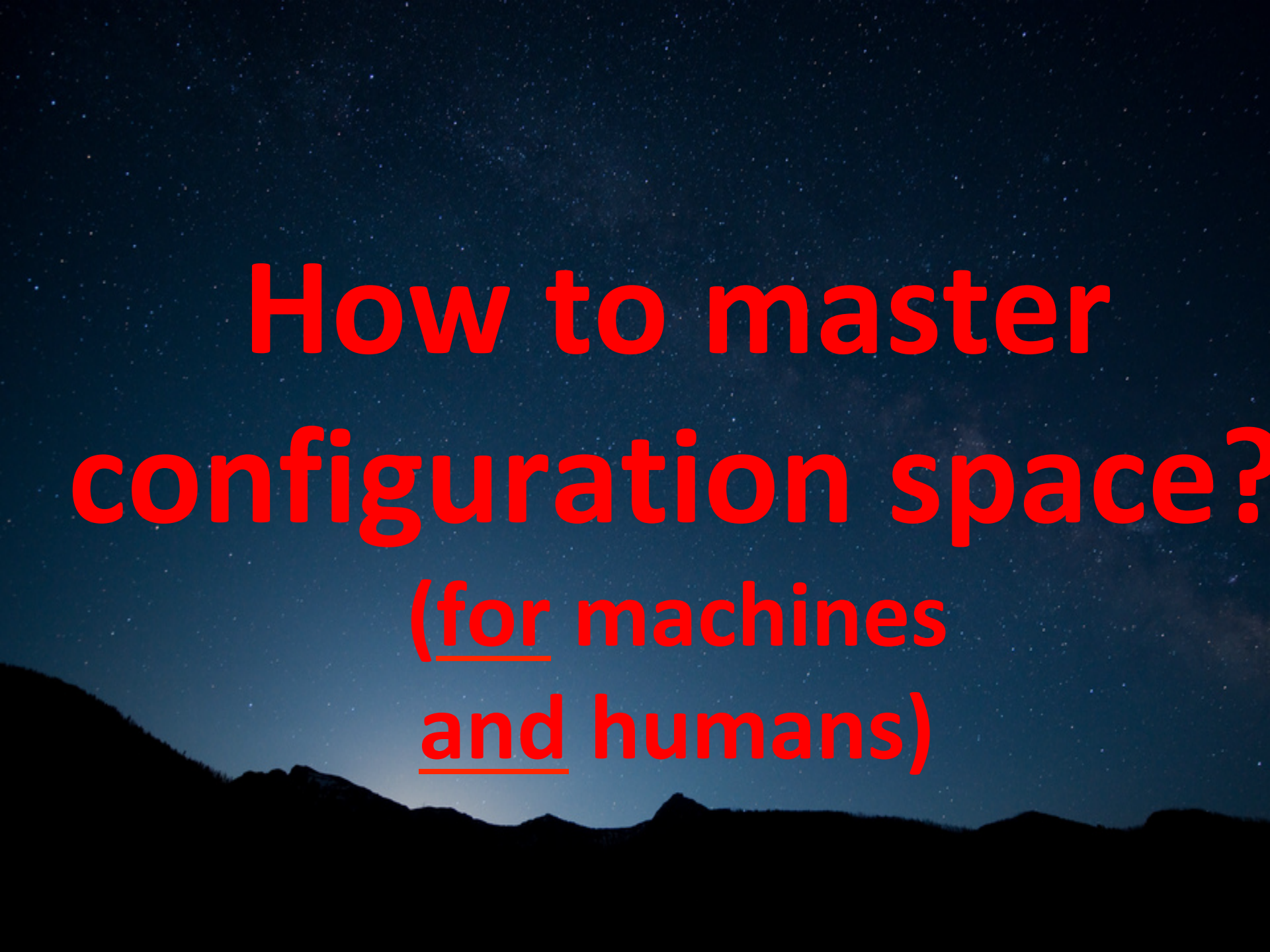


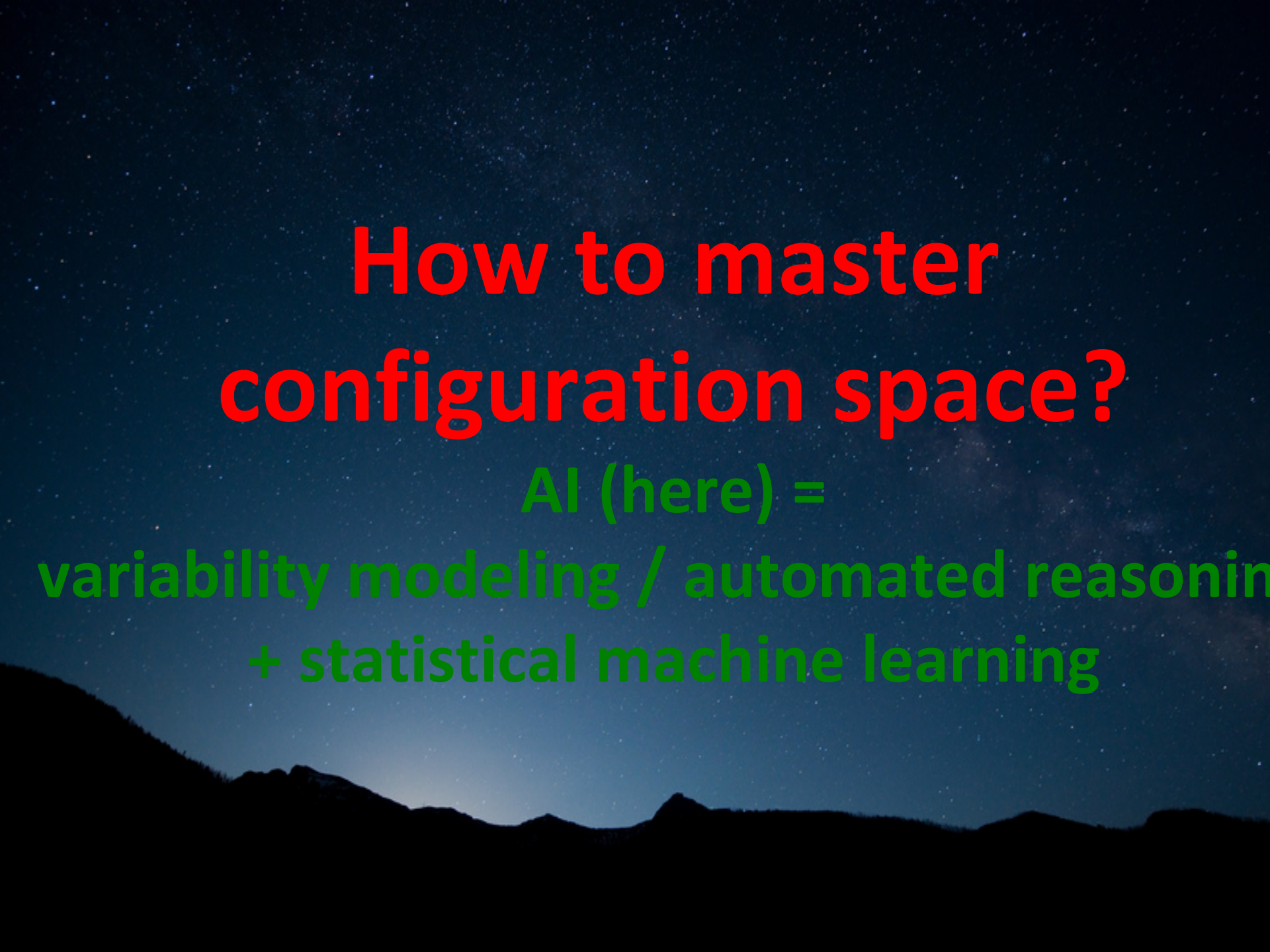
2^{33}

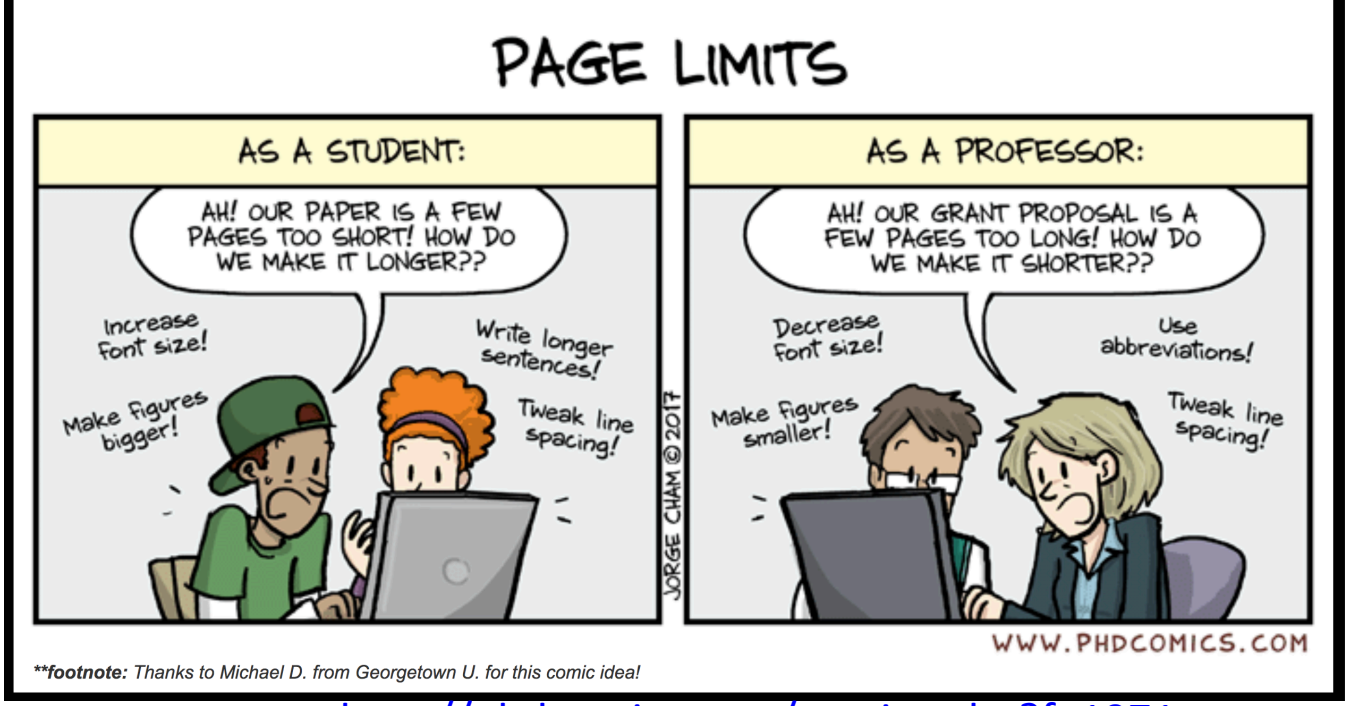
2^{320}

 # Variants (independent Boolean options)


How to master configuration space?
(with machines and humans)


How to master configuration space?
(for machines and humans)


How to master configuration space?
AI (here) =
variability modeling / automated reasonin
+ statistical machine learning



http://phdcomics.com/comics.php?f=1971

VaryLaTeX: Learning Paper Variants That Meet Constraints

Mathieu Acher
Paul Temple
Jean-Marc Jézéquel
Univ Rennes, Inria, CNRS, IRISA
Rennes, France
mathieu.acher@irisa.fr

José A. Galindo
University of Sevilla
Sevilla, Spain
jagalindo@us.es

Jabier Martinez
Tewfik Ziadi
Sorbonne University UPMC
Paris, France
jabier.martinez@lip6.fr

Successfully submitted for VaMoS'18

(on time and meeting formatting instructions)

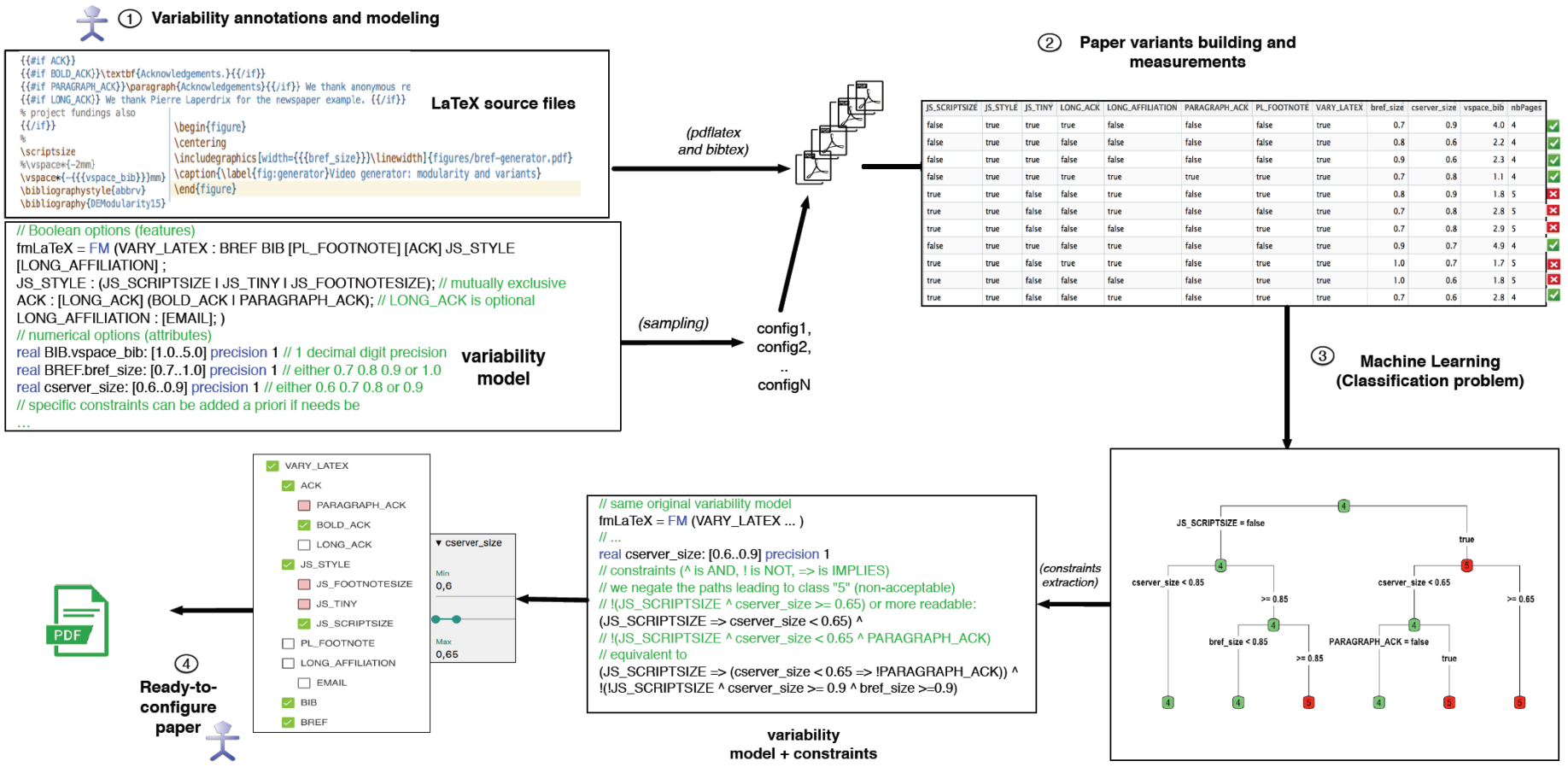
and then accepted

(live demonstration)

Two case studies

- FSE paper (see demonstration)
  - Page limit: 4
  - Accuracy: ~85% with 40 papers in the training set (there are 73,440 valid configurations)
- Curiculum vitae generation
  - 18 pages limit; 5 Boolean options; full generation, only 32 papers (not need to learn here)

Process


① Variability annotations and modeling
LaTeX source files
{{#if ACK}}
{{#if BOLD_ACK}}\textbf{Acknowledgements.}{{/if}}
{{#if PARAGRAPH_ACK}}\paragraph{Acknowledgements}{{/if}} We thank anonymous re
{{#if LONG_ACK}} We thank Pierre Laperdrix for the newspaper example. {{/if}}
% project fundings also
{{/if}}
%
\scriptsize
%\vspace*{-2mm}
\vspace*{-{{{vspace_bib}}}mm}
\bibliographystyle{abbrv}
\bibliography{DEModularity15}
\begin{figure}
\centering
\includegraphics[width={{{bref_size}}}\linewidth]{figures/bref-generator.pdf}
\caption{\label{fig:generator}Video generator: modularity and variants}
\end{figure}
(pdflatex and bibtex)
② Paper variants building and measurements
JS_SCRIPTSIZE | JS_STYLE | JS_TINY | LONG_ACK | LONG_AFFILIATION | PARAGRAPH_ACK | PL_FOOTNOTE | VARY_LATEX | bref_size | cserver_size | vspace_bib | nbPages
false | true | true | true | false | false | false | true | 0.7 | 0.9 | 4.0 | 4
false | true | true | false | false | false | false | true | 0.8 | 0.6 | 2.2 | 4
false | true | true | false | false | false | false | true | 0.9 | 0.6 | 2.3 | 4
false | true | true | true | true | true | true | true | 0.7 | 0.8 | 1.1 | 4
true | true | false | false | true | false | true | true | 0.8 | 0.9 | 1.8 | 5
true | true | false | false | true | false | false | true | 0.7 | 0.8 | 2.8 | 5
true | true | false | false | false | false | true | true | 0.7 | 0.8 | 2.9 | 5
false | true | true | false | true | false | false | true | 0.9 | 0.7 | 4.9 | 4
true | true | false | true | true | false | true | true | 1.0 | 0.7 | 1.7 | 5
true | true | false | false | false | false | true | true | 1.0 | 0.6 | 1.8 | 5
true | true | false | false | true | false | true | true | 0.7 | 0.6 | 2.8 | 4
// Boolean options (features)
fmLaTeX = FM (VARY_LATEX : BREF BIB [PL_FOOTNOTE] [ACK] JS_STYLE [LONG_AFFILIATION] ;
JS_STYLE : (JS_SCRIPTSIZE I JS_TINY I JS_FOOTNOTESIZE); // mutually exclusive
ACK : [LONG_ACK] (BOLD_ACK I PARAGRAPH_ACK); // LONG_ACK is optional
LONG_AFFILIATION : [EMAIL];)
// numerical options (attributes)
real BIB.vspace_bib: [1.0..5.0] precision 1 // 1 decimal digit precision
real BREF.bref_size: [0.7..1.0] precision 1 // either 0.7 0.8 0.9 or 1.0
real cserver_size: [0.6..0.9] precision 1 // either 0.6 0.7 0.8 or 0.9
// specific constraints can be added a priori if needs be
...
variability model
(sampling)
config1,
config2,
..
configN
③ Machine Learning (Classification problem)
JS_SCRIPTSIZE = false
true
cserver_size < 0.85
>= 0.85
cserver_size < 0.65
>= 0.65
bref_size < 0.85
>= 0.85
PARAGRAPH_ACK = false
true
(constraints extraction)
// same original variability model
fmLaTeX = FM (VARY_LATEX ...)
// ...
real cserver_size: [0.6..0.9] precision 1
// constraints (^ is AND, ! is NOT, => is IMPLIES)
// we negate the paths leading to class "5" (non-acceptable)
// !(JS_SCRIPTSIZE ^ cserver_size >= 0.65) or more readable:
(JS_SCRIPTSIZE => cserver_size < 0.65) ^
// !(JS_SCRIPTSIZE ^ cserver_size < 0.65 ^ PARAGRAPH_ACK)
// equivalent to
(JS_SCRIPTSIZE => (cserver_size < 0.65 => !PARAGRAPH_ACK)) ^
!(!JS_SCRIPTSIZE ^ cserver_size >= 0.9 ^ bref_size >=0.9)
variability model + constraints
VARY_LATEX
ACK
PARAGRAPH_ACK
BOLD_ACK
LONG_ACK
JS_STYLE
JS_FOOTNOTESIZE
JS_TINY
JS_SCRIPTSIZE
PL_FOOTNOTE
LONG_AFFILIATION
EMAIL
BIB
BREF
▾ cserver_size
Min
0,6
Max
0,65
PDF
④ Ready-to-configure paper

AI#1 Logic, satisfiability, constraints, reasoning, solving

① **Variability annotations and modeling**

LaTeX source files

```
{{#if ACK}}
{{#if BOLD_ACK}}\textbf{Acknowledgements.}{{/if}}
{{#if PARAGRAPH_ACK}}\paragraph{Acknowledgements}{{/if}} We thank anonymous re
{{#if LONG_ACK}} We thank Pierre Laperdrix for the newspaper example. {{/if}}
% project fundings also
{{/if}}
%
\scriptsize
%\vspace*{-2mm}
\vspace*{-{{{vspace_bib}}}mm}
\bibliographystyle{abbrv}
\bibliography{DEModularity15}
```

```
\begin{figure}
\centering
\includegraphics[width={{{bref_size}}}\linewidth]{figures/bref-generator.pdf}
\caption{\label{fig:generator}Video generator: modularity and variants}
\end{figure}
```

variability model

```
// Boolean options (features)
fmLaTeX = FM (VARY_LATEX : BREF BIB [PL_FOOTNOTE] [ACK] JS_STYLE
[LONG_AFFILIATION] ;
JS_STYLE : (JS_SCRIPTSIZE | JS_TINY | JS_FOOTNOTESIZE); // mutually exclusive
ACK : [LONG_ACK] (BOLD_ACK | PARAGRAPH_ACK); // LONG_ACK is optional
LONG_AFFILIATION : [EMAIL]; )
// numerical options (attributes)
real BIB.vspace_bib: [1.0..5.0] precision 1 // 1 decimal digit precision
real BREF.bref_size: [0.7..1.0] precision 1 // either 0.7 0.8 0.9 or 1.0
real cserver_size: [0.6..0.9] precision 1 // either 0.6 0.7 0.8 or 0.9
// specific constraints can be added a priori if needs be
…
```

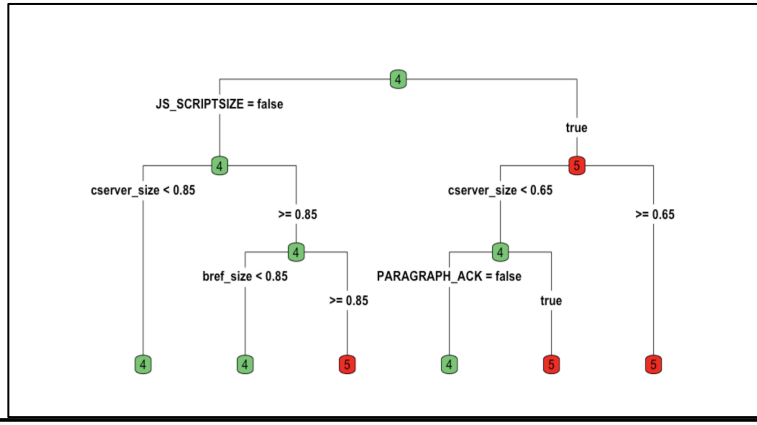
AI#2 Statistical, supervised machine learning (classification problem)

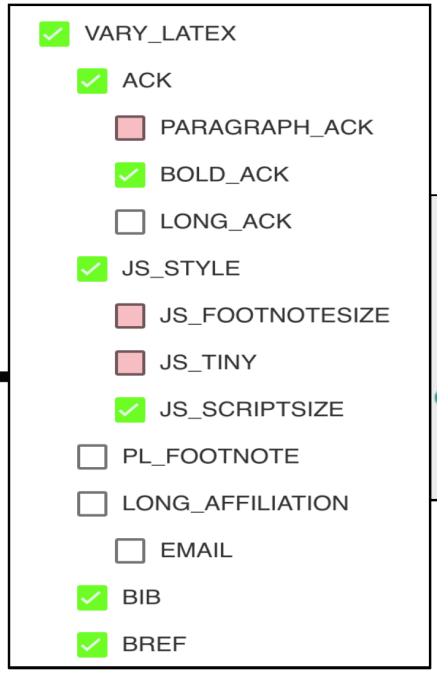
Paper variants building and measurements

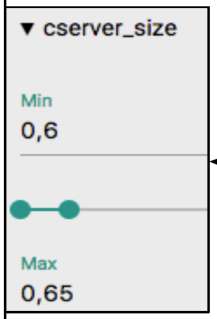
ONG_ACK	LONG_AFFILIATION	PARAGRAPH_ACK	PL_FOOTNOTE	VARY_LATEX	bref_size	cserver_size	vspace_bib	nbPages	
rue	false	false	false	true	0.7	0.9	4.0	4	✅
alse	false	false	false	true	0.8	0.6	2.2	4	✅
alse	false	false	false	true	0.9	0.6	2.3	4	✅
rue	true	true	true	true	0.7	0.8	1.1	4	✅
alse	true	false	true	true	0.8	0.9	1.8	5	❌
alse	true	false	false	true	0.7	0.8	2.8	5	❌
alse	false	false	true	true	0.7	0.8	2.9	5	❌
alse	true	false	false	true	0.9	0.7	4.9	4	✅
rue	true	false	true	true	1.0	0.7	1.7	5	❌
alse	false	false	true	true	1.0	0.6	1.8	5	❌
alse	true	false	true	true	0.7	0.6	2.8	4	✅

#AI1 + #AI2

Specialization of the variability model





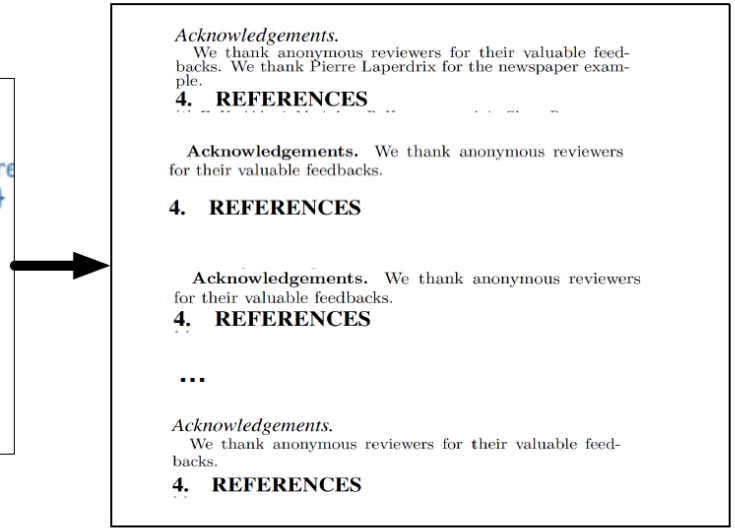



```
// same original variability model
fmLaTeX = FM (VARY_LATEX ... )
// ...
real cserver_size: [0.6..0.9] precision 1
// constraints (^ is AND, ! is NOT, => is IMPLIES)
// we negate the paths leading to class "5" (non-acceptable)
// !(JS_SCRIPTSIZE ^ cserver_size >= 0.65) or more readable:
(JS_SCRIPTSIZE => cserver_size < 0.65) ^
// !(JS_SCRIPTSIZE ^ cserver_size < 0.65 ^ PARAGRAPH_ACK)
// equivalent to
(JS_SCRIPTSIZE => (cserver_size < 0.65 => !PARAGRAPH_ACK)) ^
!(!JS_SCRIPTSIZE ^ cserver_size >= 0.9 ^ bref_size >=0.9)
```

variability

https://github.com/FAMILIAR-project/varylatex/

```
{{#if ACK}}
{{#if BOLD_ACK}}\textbf{Acknowledgements.}{{/if}}
{{#if PARAGRAPH_ACK}}\paragraph{Acknowledgements}{{/if}} We thank anonymous re
{{#if LONG_ACK}} We thank Pierre Laperdrix for the newspaper example. {{/if}}
% project fundings also
{{/if}}
%
\scriptsize
%\vspace*{-2mm}
\vspace*{-{{{vspace_bib}}}mm}
\bibliographystyle{abbrv}
\bibliography{DEModularity15}
```



Variability and LaTeX source files

Paper variants (PDF)

(a) Variability annotations and excerpt of some possible paper variants

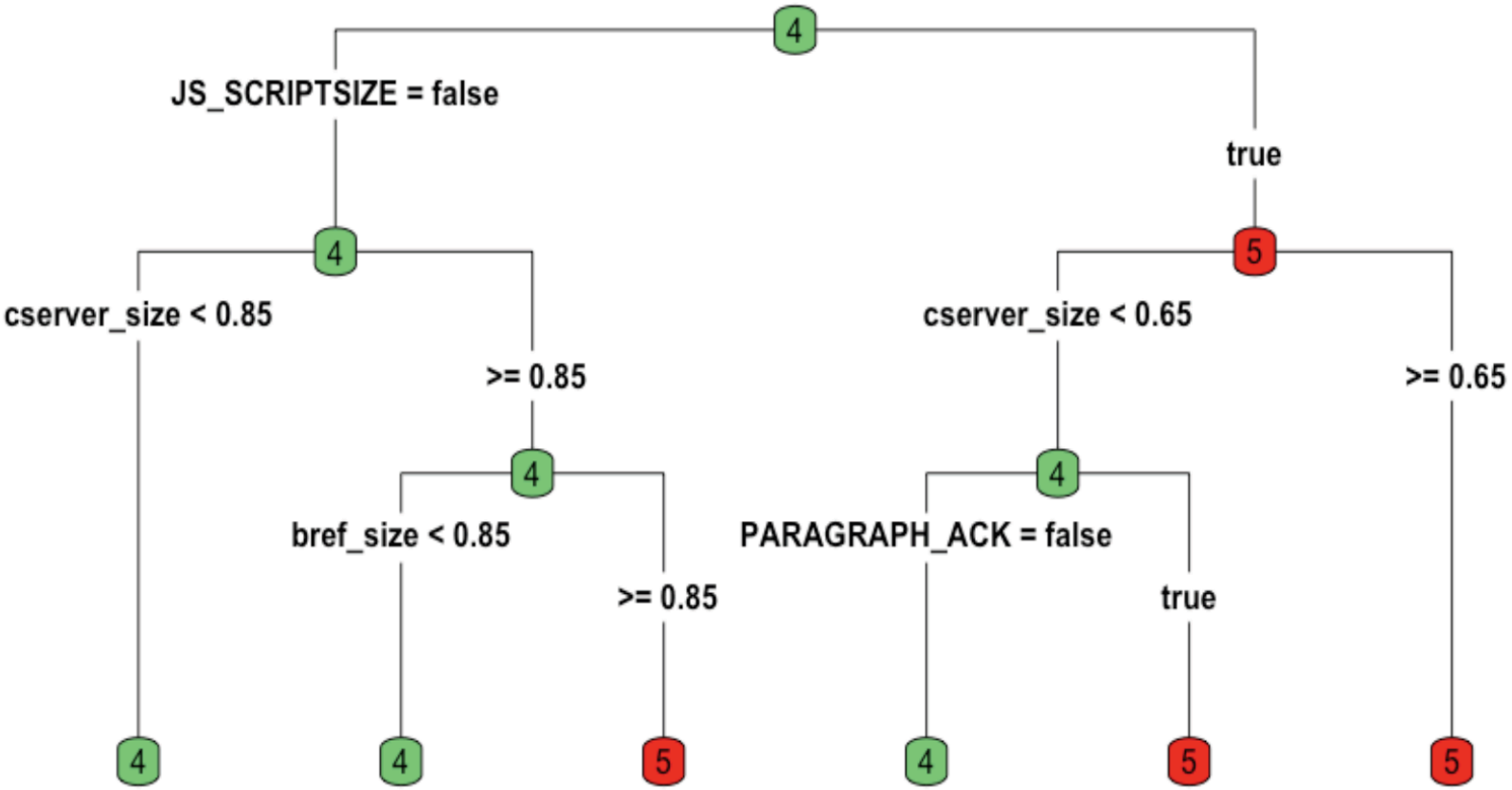
```
\lstdefinelanguage{JavaScript}{
  keywords={typeof, new, true, false, catch, function, return, null, catch, switch, var, if, in, while, do, else, case, break},
  keywordstyle=\color{blue}\bfseries,
  basicstyle=\ttfamily{{#if JS_SCRIPTSIZE}}\scriptsize{{/if}}{{#if JS_TINY}}\tiny{{/if}}{{#if JS_FOOTNOTESIZE}}\footnotesize{{/if}},
```

```
{{#if PL_FOOTNOTE}}\footnote{We are considering "product lines" in a broad sense,
```

```
\begin{figure}
\centering
\includegraphics[width={{{bref_size}}}\linewidth]{figures/bref-generator.pdf}
\caption{\label{fig:generator}Video generator: modularity and variants}
\end{figure}
```

(b) Users can vary the font size of a code snippet, activate a footnote, vary the font size of a figure, *etc.*

Classification tree

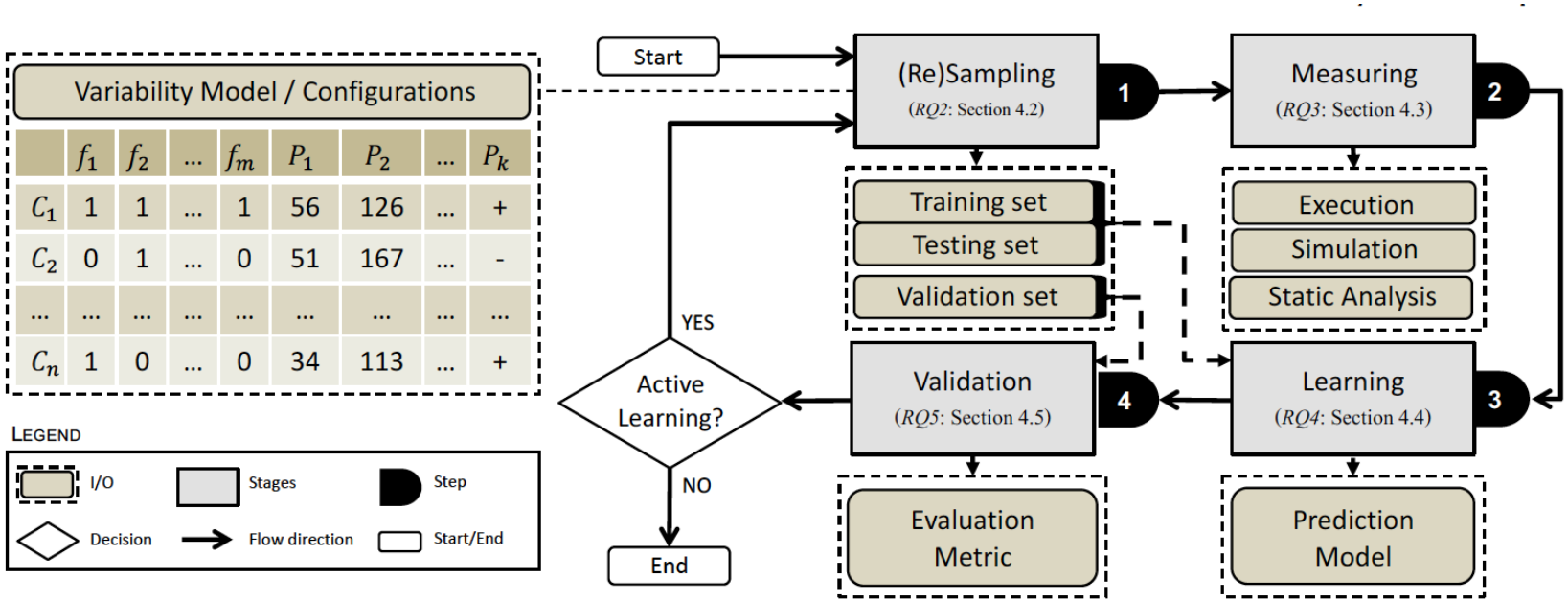


VaryLaTeX

an instance of a more general problem

(and solution based on artificial intelligence and software engineering techniques)

Sampling, Measuring, Learning



Learning Software Configuration Spaces: A Systematic Literature Review
Juliana Alves Pereira, Hugo Martin, Mathieu Acher, Jean-Marc Jézéquel, Goetz Botterweck, Anthony Ventresque https://arxiv.org/abs/1906.03018

Next slides…

- A tour of existing works
- Systematic literature survey
- Different applications (pure prediction, optimization, specialization, understanding, etc.)
- Subject systems and application domains
- Numerous sampling strategies
- How configurations are tested/measured
- Learning algorithms used and their assessment

VaryLaTeX exercice

VaryLaTeX exercice

- Instructions: https://github.com/VaryVary/ML-configurable-SPLCTutorial
- Execution of the script
- Explanations:
  - Accuracy: metrics, confusion matrix, flexibility vs safety
  - Interpretability: extraction of rules, decision tree algorithm
- Effect of
  - Training set size
  - Hyperparameters
  - Algorithms (eg random forest)

Some slides here

x264 (Hugo)

```
x264 --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

?? seconds

Guo et al. ASE 2013, Apel et al. ASE'15, Czarnecki et al. SPLC'15, Siegmund et al. FSE'15, Kastner et al. ASE'17, Menzies et al. FSE'17, Batory et al. FSE'17

```
      --psy-rd <float:float>  Strength of psychovisual optimization ["1.0:0.0"]
                                  #1: RD (requires subme>=6)
                                  #2: Trellis (requires trellis, experimental)
      --no-8x8dct             Disable adaptive spatial transform size
  -t, --trellis <integer>     Trellis RD quantization. [1]
                                  - 0: disabled
                                  - 1: enabled only on the final encode of a MB
                                  - 2: enabled on all mode decisions
      --nr <integer>          Noise reduction [0]
      --cqmfile <string>      Read custom quant matrices from a JM-compatible file

Input/Output:

  -o, --output <string>       Specify output file
      --muxer <string>        Specify output container format ["auto"]
                                  - auto, raw, mkv, flv
      --demuxer <string>      Specify input container format ["auto"]
                                  - auto, raw, y4m, avs
      --input-fmt <string>    Specify input file format (requires lavf support)
      --input-csp <string>    Specify input colorspace format for raw input
      --output-csp <string>   Specify output colorspace ["i420"]
                                  - i420, i422, i444, rgb
      --input-depth <integer> Specify input bit depth for raw input
      --input-range <string>  Specify input color range ["auto"]
                                  - auto, tv, pc
      --input-res <intxint>   Specify input resolution (width x height)
      --index <string>        Filename for input index file
      --sar width:height      Specify Sample Aspect Ratio
      --fps <float|rational>  Specify framerate
      --seek <integer>        First frame to encode
      --frames <integer>      Maximum number of frames to encode
      --level <string>        Specify level (as defined by Annex A)
      --bluray-compat         Enable compatibility hacks for Blu-ray support
      --avcintra-class <integer> Use compatibility hacks for AVC-Intra class
                                  - 50, 100, 200
      --stitchable            Don't optimize headers based on video content
                              Ensures ability to recombine a segmented encode
```

Performance Prediction

```
x264 --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

40 seconds

Performance Prediction

```
x264 --no-mbtree
     --rc-lookahead 40
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

10 seconds

Performance Prediction

```
x264 …
      -o trailer_480p24.x264
      trailer_2k_480p24.y4m
```

??? seconds

Performance Prediction

```
x264 --no-mbtree
     --rc-lookahead 40
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

??? seconds

no_8x8dct	no_asm	no_cabac	no_deblock	no_fast_pskip	no_mbtree	no_mixed_refs	no_weightb	rc_lookahead	ref	size	elapsedtime
True	False	False	True	True	False	True	True	20	9	1718492	3.444
True	False	True	False	True	False	False	True	40	9	1962957	4.744
True	False	False	True	False	True	True	False	40	1	3657562	2.427
True	False	True	False	True	True	True	False	40	9	3436410	3.447
False	False	False	True	False	False	True	False	60	5	2066645	2.957

Regression problem (linear regression, regression tree, random forest, gradient boosting, SVM, etc.)

Guo et al. ASE 2013, Apel et al. ASE'15, Czarnecki et al. SPLC'15, Siegmund et al. FSE'15, Kastner et al. ASE'17, Menzies et al. FSE'17, Batory et al. FSE'17

Classification vs regression

- Metrics
- Algorithm level (classification vs regression tree)

Exercice on x264

- MAPE metric
- Effect of
  - Training set size
  - Hyperparameters
  - Algorithms (eg gradient boosting tree)

Wrap up

Huge applicability!

Name	Domain	Non-Functional Properties
Thingiverse's	3D printer	defects
IBM WebSphere	Application server	throughput
Clasp	ASP solver	response time
SNW	Asset management	area and throughput
Binpacking	Binpacking algorithm	execution time and accuracy
XGBoost	Boosting algorithms	training time
SaaS system	Cloud computing	response time
Clustering	Clustering algorithm	execution time and accuracy
AJStats	Code analyzer	response time
SaC	Code analyzer	I/O time, response time
POLLY	Code optimizer	runtime
Libssh	Combinatorial model	defects
Telecom	Communication system	defects
LLVM	Compiler	memory footprint, performance, response time, code complexity, compilation time
Compressor SPL	Compression library	compression time, memory usage and compression ratio
7Z	Compression library	Compression time
LRZIP	Compression library	compressed size, compression time, compilation time
RAR	Compression library	code complexity
XZ	Compression library	compression time
ZipMe	Compression library	memory footprint, performance, code complexity, time
WordPress	Content management	CPU power consumption
LinkedList	Data structures	memory footprint, performance, maintainability, binary size
Curl	Data transfer	code complexity

Name	Domain	Non-Functional Properties
Wget	Data transfer	memory footprint, code complexity
Actian Vector	Database system	runtime
Apache Cassandra	Database system	latency
Berkeley DB	Database system	I/O time, memory footprint, performance, response time, code complexity, maintainability, binary size
FAME-DBMS	Database system	maintainability, binary size, performance
MySQL	Database system	defects, throughput, latency
Postgres	Database system	throughput, latency
Prevayler	Database system	memory footprint, performance
SQLite	Database system	memory footprint, performance, response time, code complexity, runtime
StockOnline	Database system	response time
Kafka	Distributed systems	throughput
DNN	DNNs algorithms	accuracy of predictions
Curriculum vitae	Document	number of pages
Paper	Document	number of pages
RUBiS	E-commerce application	response time
EMAIL	E-mail client	time
MBED TLS	Encryption library	response time
SAP ERP	Enterprise Application	response time
noc-CM-log	FPGA	CPU power consumption, runtime
sort-256	FPGA	area, throughput
E-Health System	Health	response time
$HIPA^{cc}$	Image processing	response time
Disparity SPL	Image processing	energy consumption
PKJab	Instant messenger	memory footprint, performance
IBM ILOG CPLEX	Integer solver	runtime
SPECjjbb2005	Java Server	response time, throughput
WEKA	Learning algorithm	accuracy of predictions
SVD	Linear algebra	execution time and accuracy
Trimesh	Mesh solver	iterations, response time
MBENCH	Micro benchmark	time
ACE+TAO system	Middleware software	defects
SensorNetwork	Network simulator	memory footprint, performance
Simonstrator	Network simulator	latency
NoC	Network-based system	energy and runtime
Helmholtz 3D	Numerical analysis	execution time and accuracy
Poisson 2D	Numerical analysis	execution time and accuracy
Linux kernel	Operating system	memory footprint, performance
DNN	Optimization algorithm	response time

Learning Software Configuration Spaces: A Systematic Literature Review
Juliana Alves Pereira, Hugo Martin, Mathieu Acher, Jean-Marc Jézéquel, Goetz Botterweck, Anthony Ventresque https://arxiv.org/abs/1906.03018

Open issues (Juliana slide, RQ6 of the survey)

Advanced topics

Learning the Configuration Space of Linux

Enormous configurations space eg Linux has 15K+ options, tri-state values {y, n, m}. A build takes 15 minutes on average on a recent machine

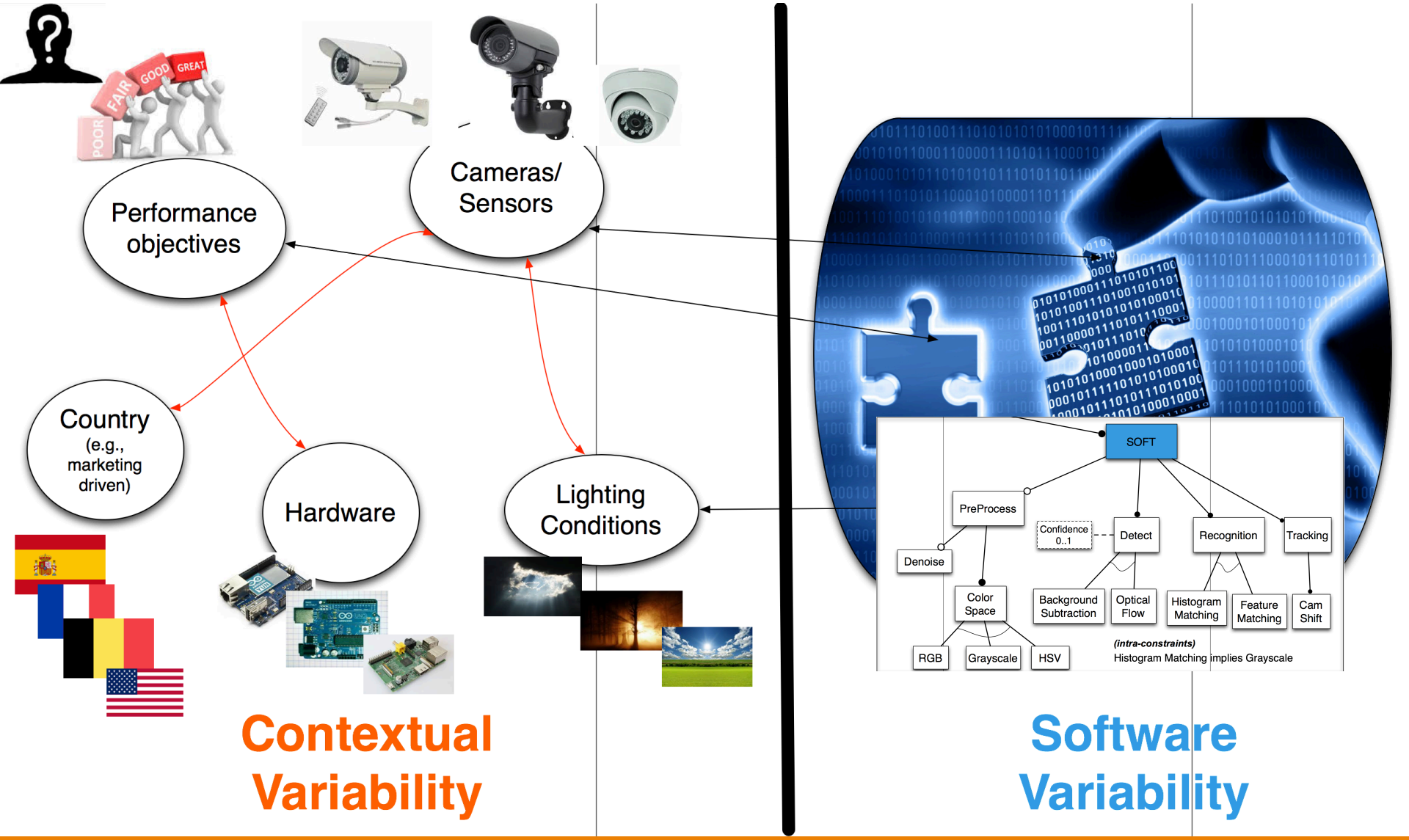
KConfig file

```
[...]

config PRINTK
	default y
	bool "Enable support for printk" if EXPERT
	select IRQ_WORK
	help
	  This option enables normal printk support. Removing it
	  eliminates most of the message strings from the kernel image
	  and makes the kernel more or less silent. As this makes it
	  very difficult to diagnose system problems, saying N here is
	  strongly discouraged.

config PRINTK_NMI
	def_bool y
	depends on PRINTK
	depends on HAVE_NMI

config BUG
	bool "BUG() support" if EXPERT
	default y
	help
	  Disabling this option eliminates support for BUG and WARN, reducing
	  the size of your kernel image and potentially quietly ignoring
	  numerous fatal conditions. You should only consider disabling this
	  option for embedded systems with no facilities for reporting errors.
	  Just say Y.

config ELF_CORE
	depends on COREDUMP
	default y
	bool "Enable ELF core dumps" if EXPERT
	help
	  Enable support for generating core dumps. Disabling saves about 4k.

[...]

config AIO
	bool "Enable AIO support" if EXPERT
	default y
	help
	  This option enables POSIX asynchronous I/O which may by used
	  by some high performance threaded applications. Disabling
	  this option saves about 7k.

[...]
```

Configurator

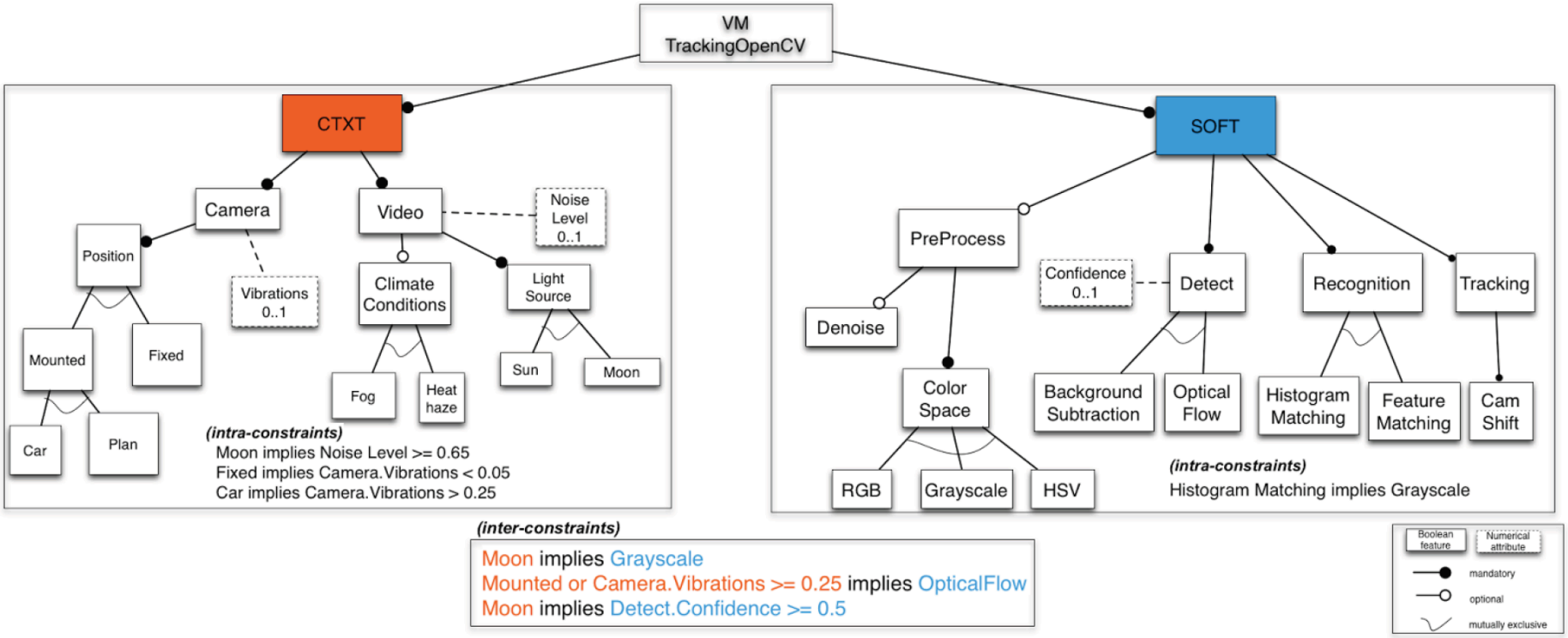
```
.config - Linux/x86 4.2.0 Kernel Configuration
                Linux/x86 4.2.0 Kernel Configuration
  Arrow keys navigate the menu.  <Enter> selects submenus ---> (or empty submenus
  ----).  Highlighted letters are hotkeys.  Pressing <Y> includes, <N> excludes, <M>
  modularizes features.  Press <Esc><Esc> to exit, <?> for Help, </> for Search.
  Legend: [*] built-in  [ ] excluded  <M> module  < > module capable

      [*] 64-bit kernel
          General setup  --->
      [*] Enable loadable module support  --->
      [*] Enable the block layer  --->
          Processor type and features  --->
          Power management and ACPI options  --->
          Bus options (PCI etc.)  --->
          Executable file formats / Emulations  --->
      [*] Networking support  --->
          Device Drivers  --->
          Firmware Drivers  --->
          File systems  --->
          Kernel hacking  --->
          Security options  --->
      -*- Cryptographic API  --->
      [*] Virtualization  --->
          Library routines  --->

          <Select>    < Exit >    < Help >    < Save >    < Load >
```

Context and Variability

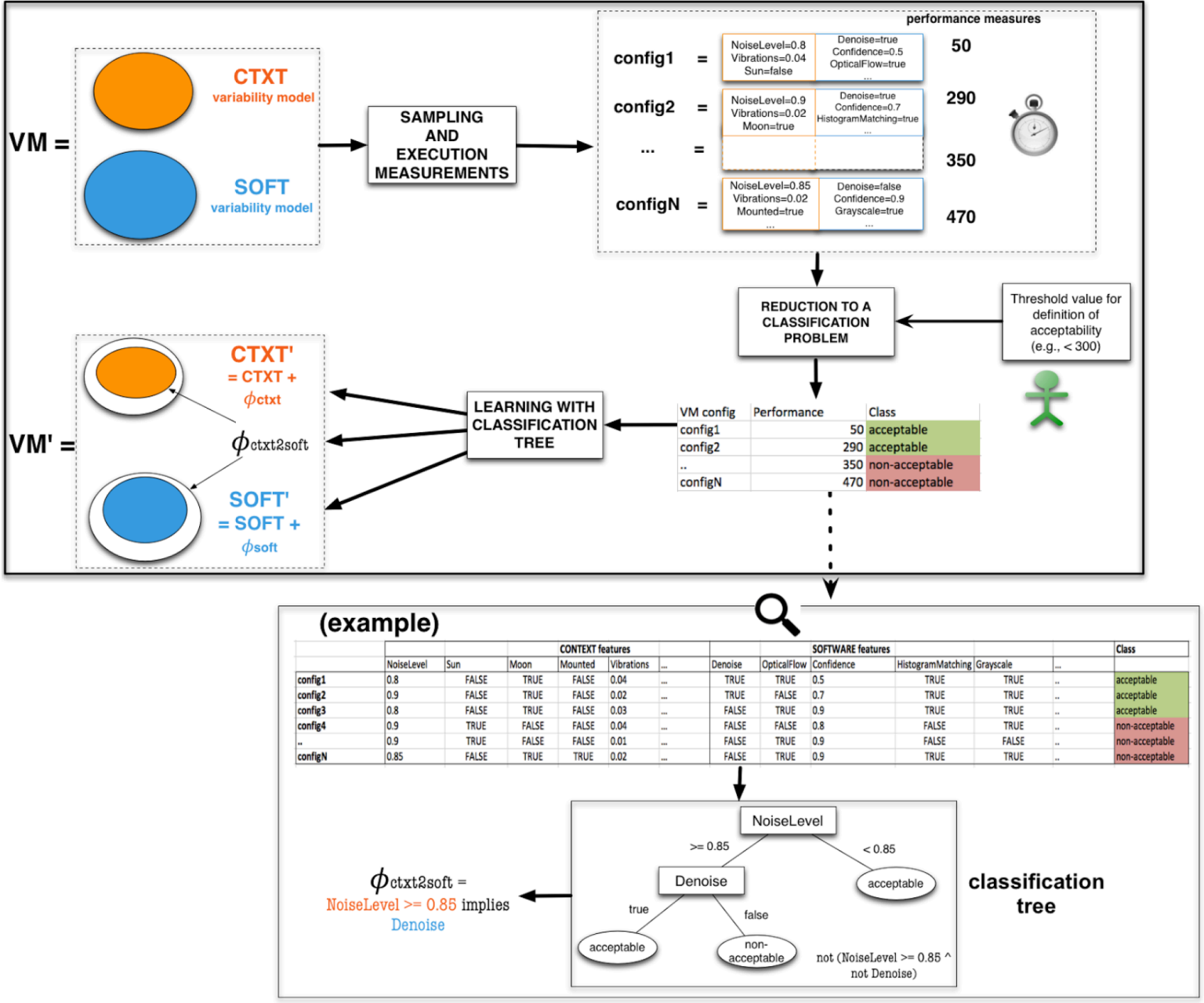


Paul Temple, Mathieu Acher, Jean-Marc Jézéquel, and Olivier Barais. Learning-Contextual Variability Models (2017). In IEEE Software

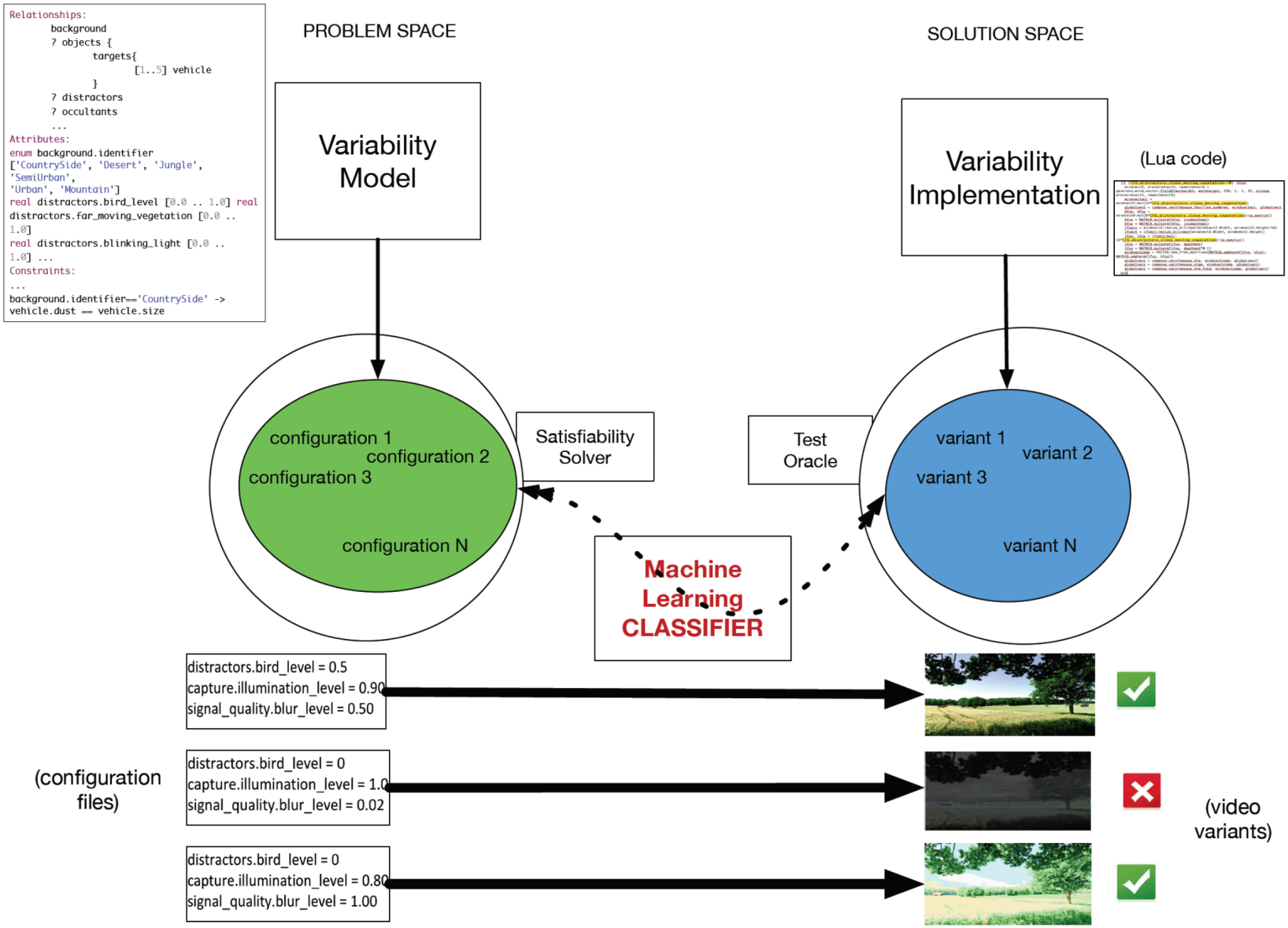
Learning
Contextual Variability Models



Paul Temple, Mathieu Acher, Jean-Marc Jézéquel, and Olivier Barais. Learning-Contextual Variability Models (2017). In IEEE Software



Paul Temple, Mathieu Acher, Jean-Marc Jézéquel, and Olivier Barais. Learning-Contextual Variability Models (2017). In IEEE Software



Paul Temple, Mathieu Acher, Battista Biggio, Jean-Marc Jézéquel, Fabio Roli:
Towards Adversarial Configurations for Software Product Lines. CoRR abs/1805.12021 (2018)

Input Sensitivity and Transferability of Performance Prediction Models

(ongoing work)

What if I change the input video?

Can I reuse my performance prediction model?

```
x264 --no-mbtree
     --rc-lookahead 40
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

```
x264 --no-mbtree
     --rc-lookahead 40
     --ref 9
     -o football.x264
     football.y4m
```

55 seconds

?

?? seconds