

# Decision Table Analysis in DMN

Presented at  
Decision Camp 2017

Bruce Silver

*bruce@brsilver.com | methodandstyle.com*

# DMN “Style”

---

- DMN is a tool-independent, graphically-defined executable decision language
- The DMN spec is quite strict, about...
  - The allowed formats of decision tables and other “boxed expressions”
  - Syntax of the FEEL expression language
  - And associated execution semantics
- ... but it allows decision tables that violate “best practices” recommended by decision modeling experts
  - ... mostly predating DMN
- Why?
  - The experts do not always agree on best practice
  - Their concepts don’t always align with DMN
  - Their requirements are not precisely defined (no formal “decision language”)
- So... “best practice” is a matter of *practitioner methodology and style*, outside the purview of the spec

# Decision Table Analysis

- Automated verification that DMN decision tables conform to these “best practices”
  - Presents errors and warnings to the modeler
  - Integrated via XML with Trisotech DMN Modeler
  - Distinct from Validation of model syntax



	Credit score	DTI	Loan Approval
U	Number	Number	tLoanApproval ["Approved", "Declined"]
1	>660	<=0.35	"Approved"
2	>=660	>=0.38	"Declined"
3	[600..660]	<=0.35	"Approved"
4	[600..660]	>0.35	"Declined"
5	<600	-	"Declined"

## Log Viewer

Severity Message

- Overlap: Rules 2, 4 match with input values [660, >=0.38]. Hit policy U allows only 1 matching rule.
- Gap: No rule matches for input values [>660, (0.35..0.38)].
- Warning: Table is not fully contracted. Combine rules 1, 3.

# Decision Tables in DMN

---

- Ultra-simple... by design
  - Condition cells (“input entries”) must be “unary tests”
    - A literal value
    - A name (of an input variable)
    - An interval between 2 of the above
    - A list of above
    - not() any of the above
    - Hyphen, meaning this input irrelevant in the rule
  - ***That’s it! No more...***
    - Arithmetic expressions not allowed
    - Disjunction, conjunction not allowed
    - Functions not allowed
  
- Why?
  1. So business users understand the logic (without training)
  2. So tool implementations are consistent/interchangeable (*good luck with that!*)
  3. To allow static analysis (of DMN style)

# What Decision Table Analysis Looks For

---

- Table completeness
- Rule overlaps and hit policy
- Masked rules
- Subsumption
- Normalization
  - First Normal Form violations
  - Second Normal Form violations
  - Third Normal Form violations

# Table Completeness

---

- Complete means every possible combination of input values is matched by a rule
- DMN1.0 had attribute *isComplete*, eliminated in DMN1.1
  - Ambiguous
    - Does it mean “intentionally” incomplete? (Do not validate completeness)
    - What about table with default output entry? Is that complete or incomplete?
- Some experts *like* incomplete tables
  - If some input value combination cannot occur, why add rule to handle it?
- Completeness detection – report gaps in the rules
  - Pre-DMN – CODASYL report 1982
  - Signavio first DMN tool to do this
  - Calvanese et al (2016), <https://kodu.ut.ee/~dumas/pubs/bpm2016DMN.pdf>
    - Algorithm minimizes count of reported gaps
  - My algorithm generates 2 test values for each condition cell “endpoint”
    - E.g. >4 generates 4 and 4+epsilon, where epsilon is a very small positive number
    - Look for test value combinations with no rule match
      - ... then convert test value back to unary test, e.g. 4+epsilon => >4
    - For performance, skip tables with >1000 possible input value combinations

# Table Completeness


Incomplete

U	Risk Score	Risk Category
	Number	<i>tRiskCategory</i> ["High", "Medium", "Low"]
1	<10	"Low"
2	(10..30]	"Medium"
3	>30	"High"

Complete

U	Risk Score	Risk Category
	Number	<i>tRiskCategory</i> ["High", "Medium", "Low"]
1	<10	"Low"
2	[10..30]	"Medium"
3	>30	"High"

## Log Viewer

Severity	Message	Element	Element Type	Page
	Gap: No rule matches for input values [10].	Risk Category	Decision Table	Risk Category

# Rule Overlaps and Hit Policy

- Rule overlap means some combination of input values matched by multiple rules
  - In that case, table outcome determined by *hit policy*
  - Unique* – Rules may not overlap
  - Any* – Overlapping rules must have same outcome (“consistent”)
  - Priority* – Select matching rule with highest outcome priority
  - First* – Select first matching rule (not declarative)
- My algorithm uses same test value generation and reporting as for gaps
  - Reports U tables with overlaps
  - ... and A tables with inconsistent outcome

Inconsistent

A	Risk Score	Risk Category
	Number	tRiskCategory ["High", "Medium", "Low"]
1	<10	"Low"
2	[10..30]	"Medium"
3	>=30	"High"

Consistent

U	Risk Score	Risk Category
	Number	tRiskCategory ["High", "Medium", "Low"]
1	<10	"Low"
2	[10..30]	"Medium"
3	>30	"High"

Severity ◆ Message



Overlap: Rules 2, 3 match with inconsistent output value using input values [30]. Hit policy A requires all matching rules to have same output value.

◆ Element ◆

Risk Category



# Masked Rules

- In P table, rule is “masked” if matches are always lower priority than another matching rule
  - Example: Masking by “else” rule that is not lowest priority
  - This error is detected by Method and Style Decision Table Analysis

Order of output values determines their priority. “Declined” must be the last one listed, the lowest priority.

P	Age	Risk Category	is Affordable	Approval Status
	Number	tRiskCategory ["High", "Medium", "Low"]	Boolean	tApprovalStatus ["Declined", "Approved"]
1	>=18	"Medium","Low"	true	"Approved"
2	-	-	-	"Declined"

P	Age	Risk Category	is Affordable	Approval Status
	Number	tRiskCategory ["High", "Medium", "Low"]	Boolean	tApprovalStatus ["Approved", "Declined"]
1	>=18	"Medium","Low"	true	"Approved"
2	-	-	-	"Declined"

Error: Rule 1 is masked by Rule2

Correct: Rule 1 is not masked

## Log Viewer

Severity	Message
✘	Masked rule error. Rule 1 is masked by other rules. Recheck output priority.

# “Misleading” Rules

- P tables can be “misleading” when rule contains hyphen input entry (“irrelevant”) and output is not lowest priority
  - Example taken from Vanthienen
- When is the *Approval Status* “Referred”?

P	isAffordable	RiskCategory	Approval Status
	<i>Boolean</i>	<i>tRiskCategory</i> ["High", "Low", "Medium"]	<i>tApprovalStatus</i> ["Approved", "Declined", "Referred"]
1	true	"Low"	"Approved"
2	-	"Medium"	"Referred"
3	true	"High"	"Declined"
4	false	-	"Declined"

U	isAffordable	RiskCategory	Approval Status
	<i>Boolean</i>	<i>tRiskCategory</i> ["High", "Low", "Medium"]	<i>tApprovalStatus</i> ["Approved", "Referred", "Declined"]
1	true	"Low"	"Approved"
2		"Medium"	"Referred"
3		"High"	"Declined"
4	false	-	"Declined"

- NOT “when *RiskCategory* is “Medium””
  - Easier to see the condition with U table
- This test not currently part of Decision Table Analysis

# Subsumption

- Subsumption means 2 rules could be combined into one (“contracted”)

Incorrect: Table is not contracted

U	CreditScore	PTI	AutoLoanApproval
	Number	Number	tAutoLoanApproval ["Approved", "Declined"]
1	<600	-	"Declined"
2	[600..660]	>=0.35	"Declined"
3		<0.35	"Approved"
4		>=0.38	"Declined"
5	>660	[0.35..0.38]	"Declined"
6		<0.35	"Approved"

Correct: Table is properly contracted

U	CreditScore	PTI	AutoLoanApproval
	Number	Number	tAutoLoanApproval ["Approved", "Declined"]
1	<600	-	"Declined"
2	>=600	>=0.35	"Declined"
3	>=600	<0.35	"Approved"

## Log Viewer

Severity	Message	Element	Element Type	Page
✘	Warning: Table is not fully contracted. Combine rules 3, 6.	AutoLoanApproval	Decision Table	AutoLoanApproval
✘	Warning: Table is not fully contracted. Combine rules 4, 5.	AutoLoanApproval	Decision Table	AutoLoanApproval

# Normalization

---

- Decision table style based on relational database principles
  - Proposed in 1990s by Vanthienen and others
    - “There is a striking similarity between decision tables and databases. The way in which we model and represent sets of rules can benefit from the insights obtained in database design. Good decision table design and good database design are subject to similar concerns, such as consistency, non-redundancy, normalization, etc.” Jan Vanthienen
  - Rediscovered and formalized by von Halle and Goldberg in The Decision Model
    - “We believe that normalization is a cornerstone for true decision modeling. Without normalization, decision modelers can create useful diagrams, but these are not the same as delivering a more formal model. The Decision Model with normalization adds simple and practical rigor to the logic of business decisions in much the same way as the Relational Model did for data.” Barbara von Halle
- As applied to DMN
  - First Normal Form
    - Rules as rows, inputs/outputs are columns
    - Column conditions always AND’ed, never OR’ed
    - No duplicate rules
    - Row/column order does not change the result
    - Inputs must be fact types not expressions
    - Output must be a single fact type (one output column)
  - Second Normal Form
    - Input must be relevant to the rule
  - Third Normal Form
    - Inputs must be independent

# First Normal Form

---

- Rules
  - Column conditions always AND'ed, never OR'ed – enforced by DMN spec
  - Rules as rows, inputs/outputs are columns – enforced by Trisotech tool
  - Inputs must be fact types not expressions – enforced by Trisotech tool
  - Row/column order does not change the result – DT Analysis violation with hit policy F
  - No duplicate rules – DT Analysis violation
  - Output must be a single fact type – Disregard; in DMN, a variable may be a structure
- TDM First Normal Form disallows multiple output columns
  - ... ***but Method and Style allows them***
  - TDM variables (“fact types”) are always simple types, but DMN allows structured types
    - Multiple output columns = multiple facts in TDM, but one variable with multiple components in DMN

# Second Normal Form

- All populated condition cells relevant to the outcome
  - Violation when rule has non-hyphen input entries irrelevant to the outcome
  - E.g., when subsumption creates irrelevant input

U	Mortgage type	Number of units	isSecondary	Max LTV
	Text ["primary residence", "second home"]	Number [1, 2, 3, 4]	Boolean	Number
1	"primary residence"	1	true	0.95
2	"primary residence"	1	false	0.95
3	"primary residence"	2,3,4	true	0.75
4	"primary residence"	2,3,4	false	0.80
5	"second home"	-	true	0.80
6	"second home"	-	false	0.85

U	Mortgage type	Number of units	isSecondary	Max LTV
	Text ["primary residence", "second home"]	Number [1, 2, 3, 4]	Boolean	Number
1	"primary residence"	1	-	0.95
2	"primary residence"	2,3,4	true	0.75
3	"primary residence"	2,3,4	false	0.80
4	"second home"	-	true	0.80
5	"second home"	-	false	0.85

Correct: Table is properly contracted

Source: <http://tdan.com/thedecisionmodeljune2013/16953>

## Log Viewer

Severity  Message



Second normal form violation. Input 3 is irrelevant for rules [1, 2]. Combine those rules for input 3.

# Third Normal Form

- Input should not be a conclusion based on other input columns
- Violation when value of 1 input can be predicted from other inputs
  - Often many gaps accompany Third Normal Form violations

U	Number of units	Mortgage balance	Mortgage eligibility	Mortgage origination date	Mortgage Relief Eligibility
	<i>Number</i> [1, 2, 3, 4]	<i>Number</i>	<i>tMortgageeligibility</i> ["Eligible", "Ineligible"]	<i>Date</i>	<i>tMortgageReliefEligibility</i> ["Eligible", "Ineligible"]
1	1	<700000	"Eligible"	<=date("2013-01-01")	"Eligible"
2	2	<900000	"Eligible"	<=date("2013-01-01")	"Eligible"
3	3	<1000000	"Eligible"	<=date("2013-01-01")	"Eligible"
4	4	<1400000	"Eligible"	<=date("2013-01-01")	"Eligible"
5	1	>=700000	"Ineligible"	<=date("2013-01-01")	"Ineligible"
6	2	>=900000	"Ineligible"	<=date("2013-01-01")	"Ineligible"
7	3	>=1000000	"Ineligible"	<=date("2013-01-01")	"Ineligible"
8	4	>=1400000	"Ineligible"	<=date("2013-01-01")	"Ineligible"
9	-	-	-	>date("2013-01-01")	"Ineligible"

Source: <http://tdan.com/thedecisionmodeljune2013/16953>

## Log Viewer

Severity  Message



Third normal form violation. Input 3 appears to depend on other inputs.



Gap: No rule matches for input values [1, (1.0E6..1.4E6), "Eligible", <=2013-01-01].

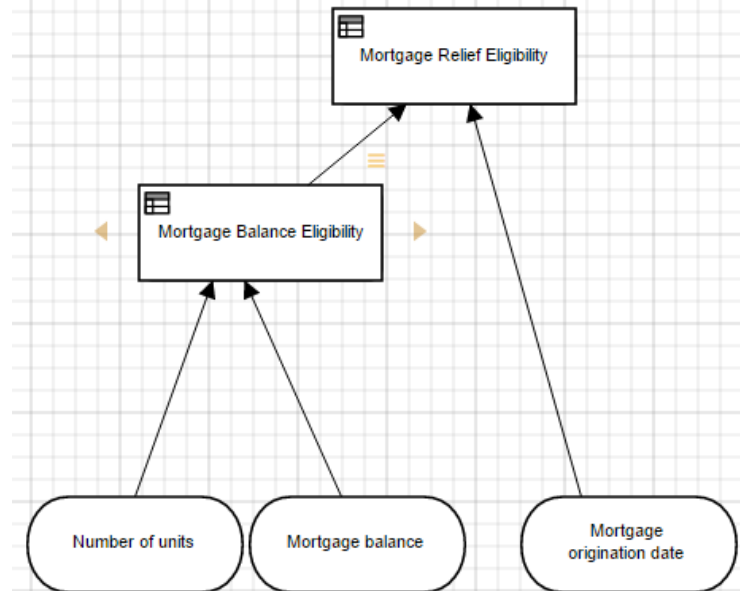


Gap: No rule matches for input values [1, (1.4E6..700000), "Ineligible", <=2013-01-01]

# Third Normal Form

- Violations fixed by breaking out dependency into a supporting decision

U	Number of units	Mortgage balance	Mortgage Balance Eligibility
	Number	Number	<i>tMortgageBalanceEligibility</i> ["Eligible", "Ineligible"]
1	1	<700000	"Eligible"
2	2	<900000	"Eligible"
3	3	<1000000	"Eligible"
4	4	<1400000	"Eligible"
5	1	>=700000	"Ineligible"
6	2	>=900000	"Ineligible"
7	3	>=1000000	"Ineligible"
8	4	>=1400000	"Ineligible"



A	Mortgage Balance Eligibility	Mortgage origination date	Mortgage Relief Eligibility
	<i>tMortgageBalanceEligibility</i> ["Eligible", "Ineligible"]	Date	<i>tMortgageReliefEligibility</i> ["Eligible", "Ineligible"]
1	"Eligible"	<=date("2013-01-01")	"Eligible"
2	"Ineligible"	-	"Ineligible"
3	-	>date("2013-01-01")	"Ineligible"



# Conclusions

---

- DMN provides a serialized language for verifying decision table conformance to “best practice” guidelines
- Errors are easy to make by accident
- Decision table analysis does not prevent execution
- Decision table analysis not only improves model quality but teaches users best practice

# For More Information

---

- Contact me: [bruce@brsilver.com](mailto:bruce@brsilver.com)
  - Twitter @bpmswatch
  - Web: [methodandstyle.com](http://methodandstyle.com)
- Register for **free trial of Trisotech DMN Modeler**
  - [www.trisotech.com/MethodandStyleTrial](http://www.trisotech.com/MethodandStyleTrial)
  - Includes **FREE** Overview + Part1 of *DMN Method and Style Basics* training
- Hands-on DMN training (web/on-demand)
  - Includes 60-day use of Trisotech DMN Modeler + **post-class certification**
  - DMN Basics: DMN concepts, DRD decomposition, decision tables, BPMN integration...
  - DMN Advanced: FEEL, BKMs, Contexts, calendar arithmetic, lists and tables, iteration, ...
  - Ridiculous bundled price through July 31
  - More info at [methodandstyle.com/product/dmn-basics-on-demand/](http://methodandstyle.com/product/dmn-basics-on-demand/)
- DMN Method and Style (book)
  - <https://www.amazon.com/DMN-Method-Style-Practitioners-Decision/dp/0982368151>

