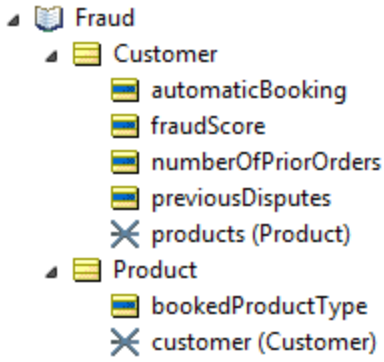
Fraud Rule Modeling Challenge Feb 2017 Corticon Solution Mike Parish

The Rules

1. If Booked Product Type is POST-PAID-HOTEL add 5 to the score
2. If Booked Product Type is INTERNAL-FLIGHT add 100 to the score
3. If Booked Product Type is INTERNATIONAL-FLIGHT add 25 to the score
4. If Booked Product Type is CAR add 10 to the score
5. If Booked Product Type is PRE-PAID-HOTEL add 5 to the score
6. If there were no previous orders from this customer add 100 to the score
7. If Number of Orders from this customer between 1 and 10 including bounds add (100 – Number of Orders * 10) to the score
8. If Customer has previous disputes add 190 to the score
9. If score is less than 200 the booking can be made automatically

The Vocabulary

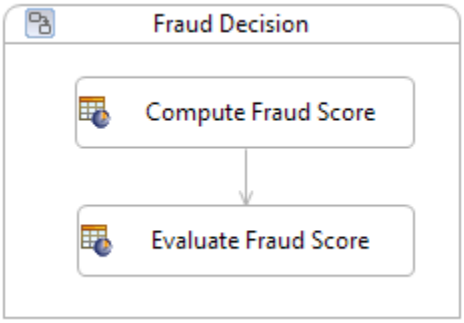
We'll use this vocabulary assuming that customers can book more than one product:



For simplicity we'll also assume that the prior order count and the previous disputes fields have already been populated, though it's relatively easy for Corticon to determine those if we had the relevant rules.

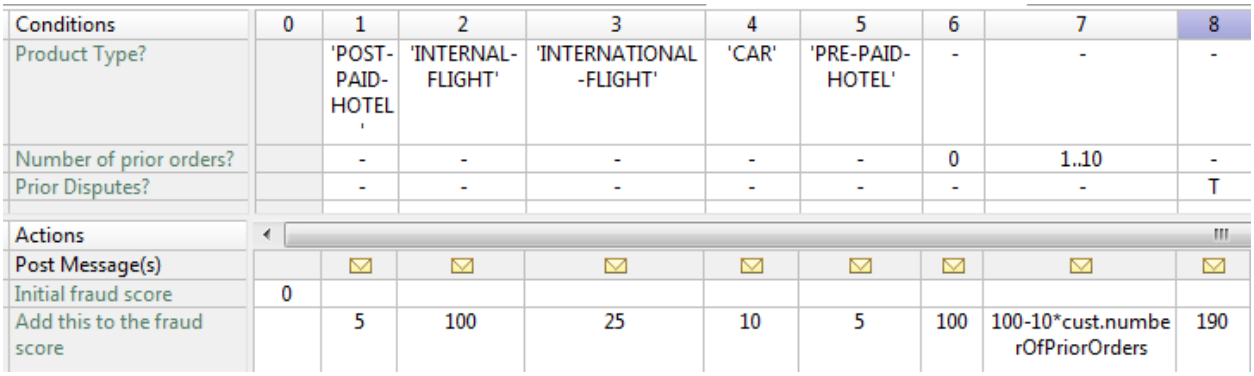
Decision Structure

The problem can be solved in two steps



Compute Fraud Score

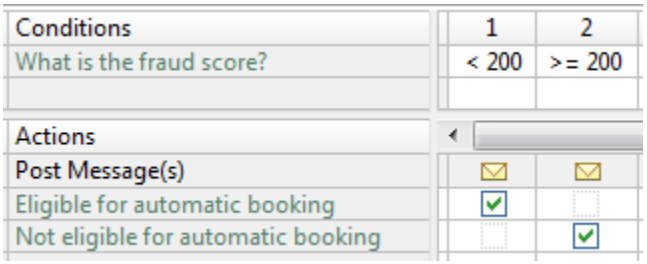
We can put the fraud scoring rules into a single decision table like this:



Conditions	0	1	2	3	4	5	6	7	8
Product Type?		'POST-PAID-HOTEL'	'INTERNAL-FLIGHT'	'INTERNATIONAL-FLIGHT'	'CAR'	'PRE-PAID-HOTEL'	-	-	-
Number of prior orders?		-	-	-	-	-	0	1..10	-
Prior Disputes?		-	-	-	-	-	-	-	T
Actions									
Post Message(s)		✉	✉	✉	✉	✉	✉	✉	✉
Initial fraud score	0								
Add this to the fraud score		5	100	25	10	5	100	100-10*cust.numberOfPriorOrders	190

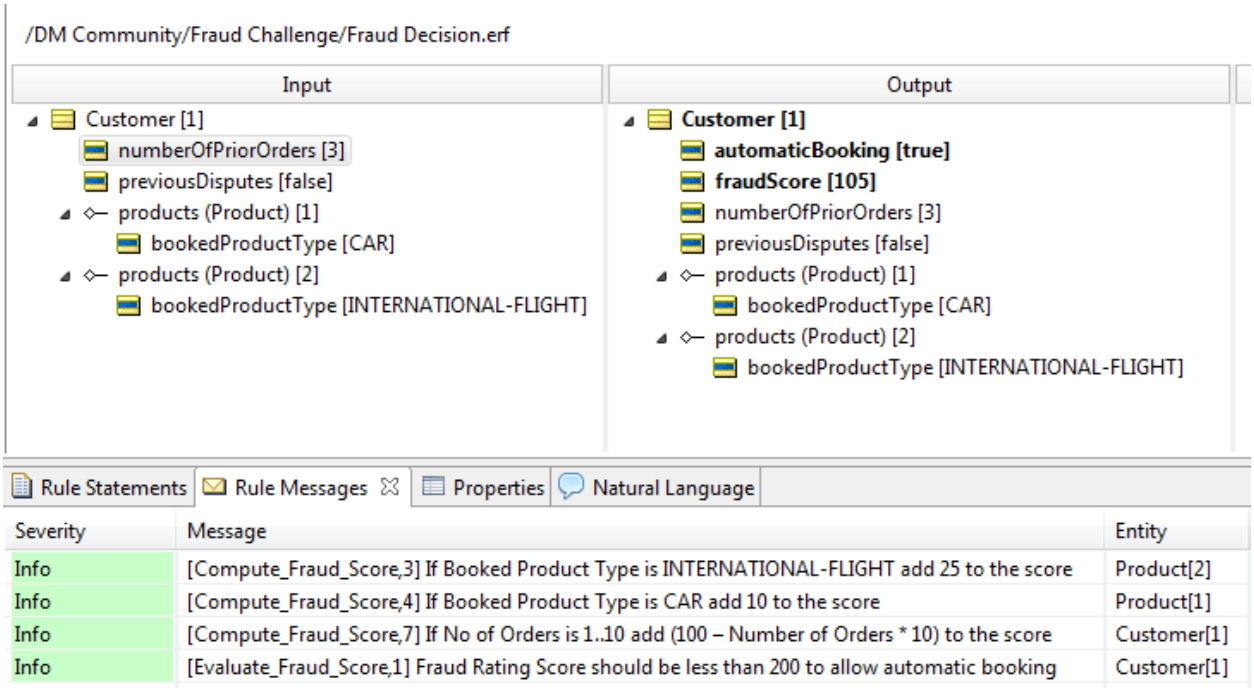
Or we could choose to split the logic into several decision tables. Since there are only eight rules currently a single table seems easier to deal with. But if the rules start to grow it may make sense to split it up. This pattern is often seen in health risk assessments where a separate rule sheet handles all the risk scores for a given risk factor. Health risk assessments may have dozens of different risk factors and hence rule sheets. See example at the end of this document.

Evaluate Fraud Score



Conditions	1	2
What is the fraud score?	< 200	>= 200
Actions		
Post Message(s)	✉	✉
Eligible for automatic booking	☑	☐
Not eligible for automatic booking	☐	☑

Test Case



Severity	Message	Entity
Info	[Compute_Fraud_Score,3] If Booked Product Type is INTERNATIONAL-FLIGHT add 25 to the score	Product[2]
Info	[Compute_Fraud_Score,4] If Booked Product Type is CAR add 10 to the score	Product[1]
Info	[Compute_Fraud_Score,7] If No of Orders is 1..10 add (100 – Number of Orders * 10) to the score	Customer[1]
Info	[Evaluate_Fraud_Score,1] Fraud Rating Score should be less than 200 to allow automatic booking	Customer[1]

Possible Enhancements

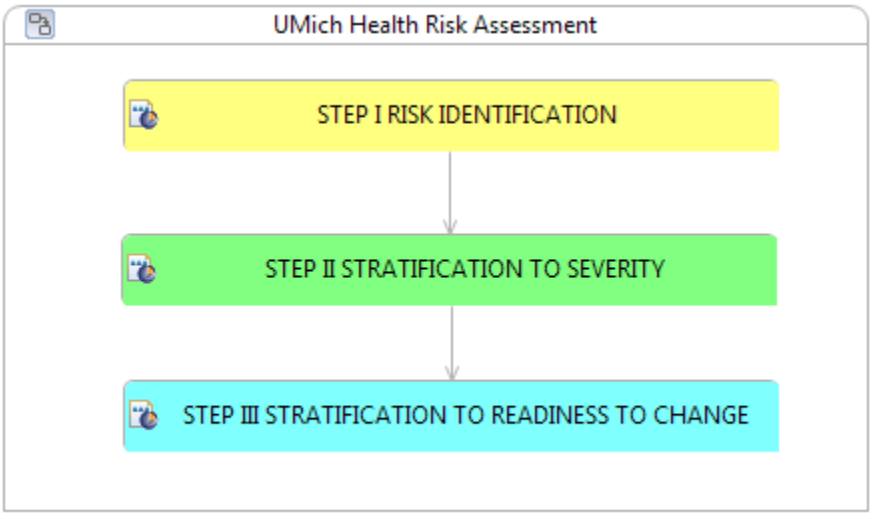
Prior Orders

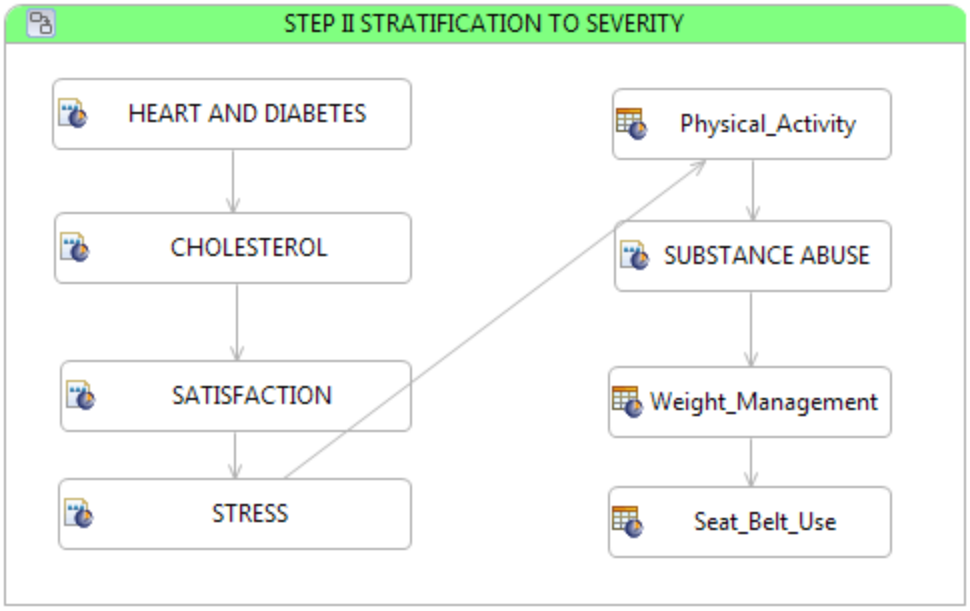
The number of prior orders could be the subject of more complex rules.
Maybe we only count orders in the past six months that are for particular products.
The more complex the logic to determine the number of orders, the more it makes sense to define the logic in a rule engine. If we simply want the total count of all orders for that customer then a simple SQL query may be more efficient.

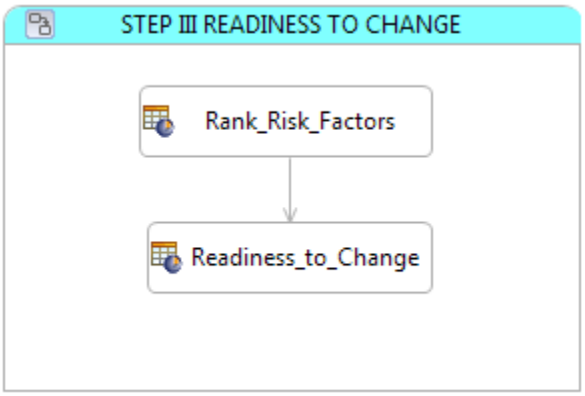
Prior Disputes

Similarly determining whether a dispute needs to be considered could involve complex logic.

Health Risk Assessment Scoring (Similar Pattern)







Examples of Health Risk Scoring Rule Sheets

Number of Risk Factors

Conditions	0	1	2	3	4	5
How many risk factors does the participant have?		>= 5	4	3	2	{0, 1}
Actions						
Post Message(s)		✉	✉	✉	✉	✉
Add this to all risk scores		10	8	6	4	2

Weight Management

Conditions	1	2	3	4	5	6	7	8	9	10
What is the participant's BMI?	26..29	30..34	35..39	>= 40	-	-	-	-	-	-
How many servings (of what during what period?)	-	-	-	-	3..4	5..6	-	-	-	-
What is the WLQ score	-	-	-	-	-	-	'hig...	-	-	-
How often does the participant engage in Brisk physical activity 20 minute sessions per week	-	-	-	-	-	-	-	>= 3	{1, 2}	0
Actions										
Post Message(s)	✉	✉	✉	✉	✉	✉	✉	✉	✉	✉
Add this to the weight management risk score	3	4	5	6	1	2	4	0	3	4

Stress

Conditions	1	2	3	4	5	6	7	8
How often is the participant Anxious, tense or depressed	'never'	'rarely'	'sometimes'	'often'	-	-	-	-
What effect is stress having on the paricipant's health	-	-	-	-	'none'	'some'	'a lot'	-
What is the WQL Score	-	-	-	-	-	-	-	'high'
Actions								
Post Message(s)	✉	✉	✉	✉	✉	✉	✉	✉
Add this to the stress risk score	0	2	4	6	0	4	6	4

A health risk assessment such as this would typically be driven by a questionnaire. Even the questions themselves might be dynamically derived by the rules depending on which risk factors need to be considered. For example if you answer “male” to the gender question, there’s no need to ask if you are pregnant. This can make a questionnaire much more palatable to the user.