Provisional Application for United States Patent

TITLE: CMO Cashflow Rule-Base Library Builder

INVENTOR(S): Weimin Zhao, John Chwanshao Wang

USPTO PATENT APPLICATION NUMBER: 62485391

BACKGROUND

A CMO (Collateralized Mortgage Obligation) is a type of mortgage backed securities that consists of multiple collateral groups, the cashflow from which are allocated to a set of tranches based on certain payment rules.

The principal payment rules to the many tranches within a collateral group determine the amount of the payment allocated to each tranche, from the principal payment received backing collateral group. The rules are described in natural language in the CMO deal's prospectus supplement. These PDF documents made available to the investment public by Fannie Mae, Freddie Mac, and Ginnie Mae.

To compute the cash flow to each tranche within a group of a CMO deal, one needs a computer program that implements the logic described by the payment rules.

BRIEF SUMMARY OF THE INVENTION

With many thousands of CMO deals issued by the agencies, where each deal has different set of tranches and associated payment rules, transforming these rules which are described in natural language inside PDF documents, and developing computer programs to implement these rules become a challenge. This patent application addresses this challenge with a novel design of methodology.

Compared to programming languages, natural languages have large selection of vocabularies and much variations in representation of any single semantic concept. In addition, it also has a large degree of freedom in expressing the logical relation among different parts of a sentence. To faithfully represent the rules expressed in the prospectus supplements, we select a widely-understood language, YAML, to express these rules, and to demonstrate our design concepts. Such a rule-base implemented in YAML contains a set of payment directives and can be easily processed by computer programs.

The fundamental approach to make as a faithful representation of the original payment rules in natural language with a set of structured payment directives is to formulate a multi-step procedure. The multi-step approach is aimed at the separation of concerns and reduction of complexity. Each step concentrates on one specific aspect of the rule text to address a specific attribute of the payment rule. At the same time, each step reduces the complexity of the problem by unifying the vocabulary of the payment rules while capturing all the critical payment information contained in the original text.

DETAILED DESCRIPTION AND BEST MODE OF IMPLEMENTATION

In what follows, we document the steps of our novel approach for the creation of such rule-base library in YAML.

1. There are some assumptions concerning CMO's principal distribution which we carry over to the YAML rule-base. As such, we can eliminate the verbiage spent on these presumptions, leading to the simplified rule text. Such assumption could include the following:

a. The rules governing the cash flow distribution only concern the current month unless specified otherwise

b. The balance with its variations in the document is the current month's unpaid principal balance

c. Each payment statement is completely contained in a sentence, unless it is a nested list.

2. Very often, different nouns are used to describe the same entity, such as principal balance, interest accrued, etc. These variations are detected using frequency analysis of the rule text and then unified by a lookup dictionary along the process. After this step, we create a limited set of vocabularies describing the entities involved in the cash flow computation as well as correspondence to the original text.

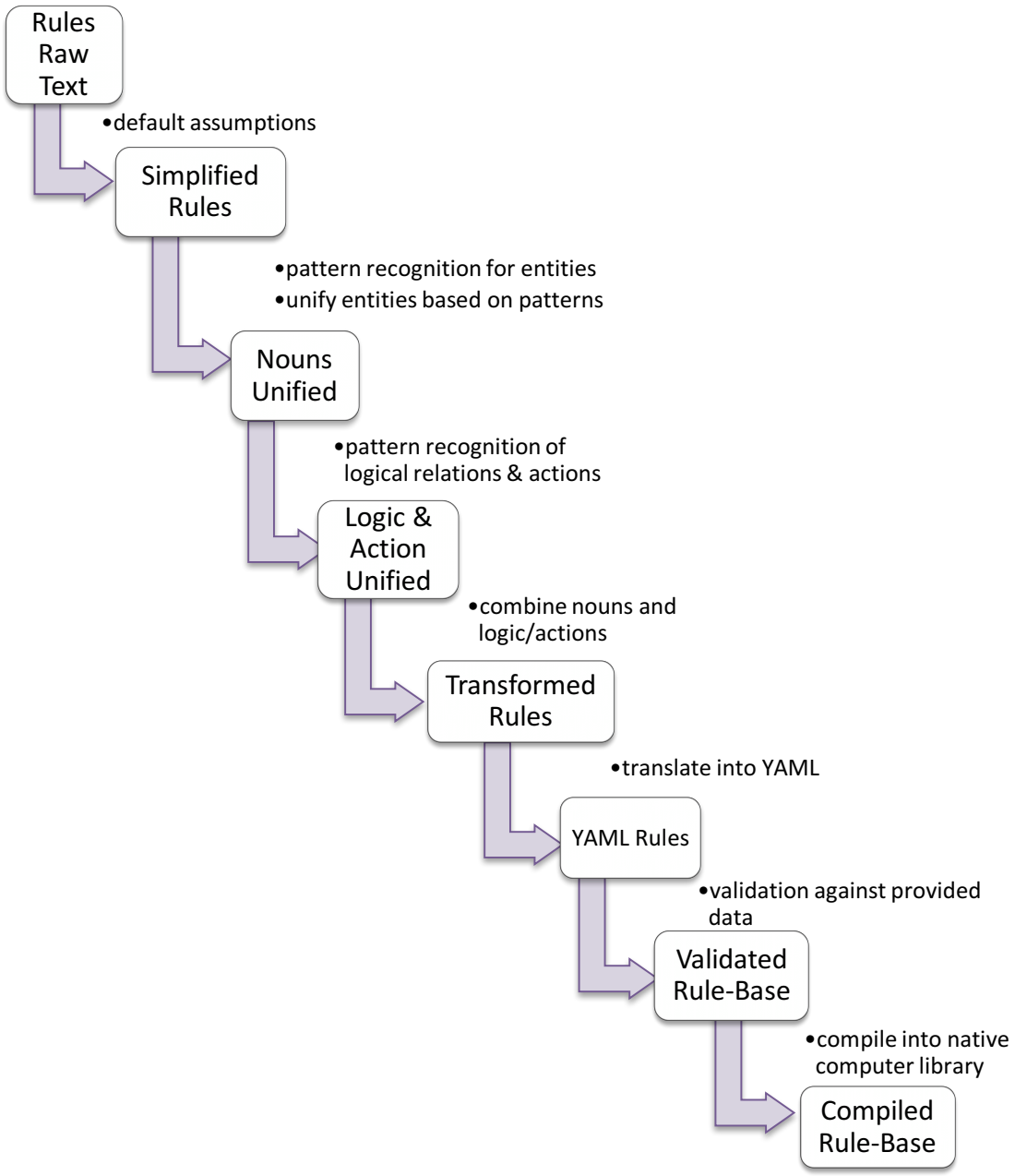

Figure 1. Seven-Step Process to Create a YAML Rule-Base for CMO Deals

3. The logical relations and actions concerning the following are unified in this step

a. The type of allocation, such as sequential allocation, pro rata allocation, etc.

b. The direction of payment, e.g., among different segments

c. Payment ending condition, e.g., until retired, until to a predefined balance, etc

d. Payment timing conditions

e. Payment dependency conditions are unified to an adaptively generated set of keywords

4. With the original text reduced to a set of keywords covering both entities and relations from earlier steps, frequency is computed on the whole sentence, with each component replaced by its functional type. Such a frequency result shows the different types of the statement structure. These patterns are then mapped back to each sentence to produce the transformed payment rule sets.

5. At this point, each payment rule described in the natural language are faithfully and completely represented by transformed rule statement, which is further translated into YAML statements.

6. To validate the rule set, the incoming collateral cash flow is computed using prepayment assumptions and mortgage amortization formula. The collateral cash flow is allocated, for each month, among the tranches within the collateral group, using the rules now expressed in YAML. It produces the time series of the principal balance and the weighted average for each tranche. This tranche balance time series and weighted average life are compared with the data listed in the balance decrement data contained in the CMO's prospectus supplement. If the exact match is not found, we test additional interpretations of the YAML statements, until exact match is found. The matching interpretation is then recorded as part of this group's YAML rule-base.

7. A compiler program compiles all group's YAML rule sets into a highly efficient cash flow rule-base library for computing CMO deals cash flow.

CLAIMS:

A method to convert principal payment rules from natural language in security prospectus supplements into a form that can be used by a computer program to calculate cash flow.

ABSTRACT

A seven-step procedure is described here to demonstrate how a YAML rule-base library can be created from the payment rules described in natural language, contained in the deal's prospectus supplement for all the CMO deals issued by the agencies. Such a library can be used in many applications where exact balance of CMO tranches are needed, given any input cash flows from the backing collaterals.