OFFICE OF FINANCIAL RESEARCH AND OPERATIONAL RISK

BY

Willi Brammertz[1]

[1] Special thanks go to Rami Entin, Allan Mendelowitz and Vicky Lemieux for their careful reading, correcting and improving the arguments presented in this paper. The presented thoughts, although theoretically well founded, are equally well founded in practice – tested in more than 200 banks in about thirty countries.

ABSTRACT

We argue that the single largest operational risk for the OFR is to be inundated with un-decipherable financial contract data. We continue to argue, that clear semantics on the attribute level alone is not sufficient to address this risk. Nevertheless, the OFR does need a semantic capable of describing the entire intent of the financial contract. Next we develop the concept of Contract Types (CT), which encapsulate this semantic. The idea of CT's requires the melding of both, data and algorithms, the separation of which lies at core of the observed data chaos in financial institutions. We distinguish the mechanistic parts of finance where the separation of data and algorithms is counter productive and the subjective parts, where the separation makes sense. We conclude with a model that makes the OFR operational.

1. ON OPERATIONAL RISK

Dodd-Frank does not mandate the OFR to control the OR of the market participants, as is the case with the Basel agreement. However, the OFR itself faces massive operational challenges and risks itself. This article addresses the OR that the OFR will have to overcome if it is to succeed and make its intended contributions to monitoring threats to financial stability.

Unlike market risk, where fluctuations are exogenous and can only be mitigated by managing exposures, the actors and "risk factors" of operational risk (OR) are the people and business processes involved. Therefore the main emphasis of OR management is the avoidance of risks in the first place. Doing otherwise would mean treating our own actions as a pure risk factor to which we are exposed similar to a foreign exchange rate risk, a proposition most people would dismiss.

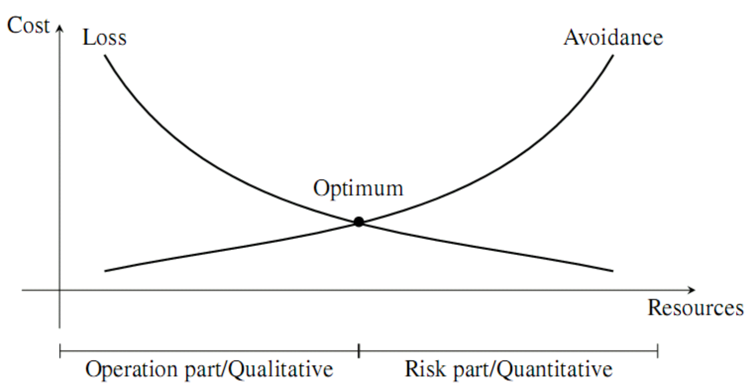


FIGURE 1: OPERATIONAL RISK

Figure 1 illustrates this relationship of risk avoidance and losses incurred. The x-axis indicates the resources allocated to the avoidance of risk and the y-axis we find the corresponding expected losses associated with OR. The "Avoidance" curve shows the implied cost related to activities that help avoid unwanted results. Initial efforts are highly effective and therefore cheap. The marginal cost of avoidance however rises as more infrequent and/or unlikely cases are dealt with. The "Loss" curve shows the marginal impact of the incurred risks. While doing nothing would be accompanied by very high cost, these marginal costs drop after initial risk avoidance initiatives have been put in place. The rate of decrease however is reduced as the more infrequent and unlikely cases are tackled.[2].

There is an optimal point where the marginal cost of additional efforts to avoid risk is equal to the marginal benefit from doing so. The region to the left of the optimum is inefficient since it is cheaper to avoid risks then to incur the corresponding losses, while the region to the right side is inefficient for the opposite reason.[3] The "Optimum" determines reasonable endeavor. These regions are often referred to as "Qualitative OR" and "Quantitative OR" or alternatively the "operation part" versus the "risk part". Only "Quantitative OR" is risk in the commonly understood sense and comparable to other types of risks such as market or credit risk.

In this paper we propose how to deal with the left side, the reduction of organizational inefficiencies to an economically efficient level. We have to ask ourselves, which are the threats to avoid with the highest possible impact. Since another characteristic of OR is its meager or even non-existent statistical base, we only can guess. The biggest threats to the OFR are probably of a political order. This risk is however the least controllable, and is in fact a real risk that resides on the right side of the optimal point. Following that, the most likely major risk is probably building a major organization, which fails to deliver on its promise due to principal flaws in the data gathering process. This is the risk we investigate in this article.

2. Tasks and risks of the OFR

In order to better understand the controllable operational risks we have to understand the tasks assigned to the OFR. The main tasks of the OFR are in short:

- Increasing the analytical capability of the government to understand and respond to threats to financial stability
- Performing independent analysis of risk
- Allow for the following (non exhaustive list)
    - Stress testing
    - Sensitivity analysis (to any risk factor)
    - Exposure analysis

These tasks are to be achieved primarily by unique data collection and analytics:

- The collection of daily data on a granular level
    - On a single contract level
    - On a single counterparty level including relationships
- Standardization of the data with
    - A Counter-party reference data base
    - A financial instrument reference data base

Another main task is the building of a multidisciplinary/interdisciplinary research staff, which can make sense of the data collected and derive – if possible before the fact - insights for the political decision makers.

[2] To be more precise: the cost of loss is defined by its impact multiplied with its frequency. If high impact low frequency cases are considered very costly then they are found on the left side of the graph, else on the right side.

[3] Of course, the cost of losses cannot be counted on a purely monetary basis but must include human hardship etc.. It has to be recognized though, that even after taking all this into account, there is a point where risks have to be accepted even if high – on a purely economic basis.

Taking into account the millions or maybe billions of financial contracts in existence at any point in time, their diversity and implied complexity, it is obviously a colossal and daring undertaking likened by some to Los Alamos or CERN. Although a flattering comparison, we think it is an overstatement, notwithstanding being a huge challenge. The magnitude of the task becomes all the more obvious if we look at the data challenges faced by banks today. Even the best organized banks are described as chaotic in terms of data and analytics, a situation that has considerably contributed to the near melt down of the financial sector in the opinion of many that should know.

This very near melt down was the impetus for the formation of the OFR. The question is then, how can the OFR avoid the same data disaster that banks have been unable to solve thus far. This disaster can only be avoided if the OFR is built on sound first principles that can be communicated, understood and applied throughout the entire financial sector. Failing that, the project will end up in an expensively produced morass. This we consider the single most important operational risk that can be solved managerially. We also believe, that the marginal cost of avoiding it is very low, especially if compared with the huge marginal benefit. In reference to Figure 1 it is a high likelihood high impact case far to the left of the optimal point.

3. First principles

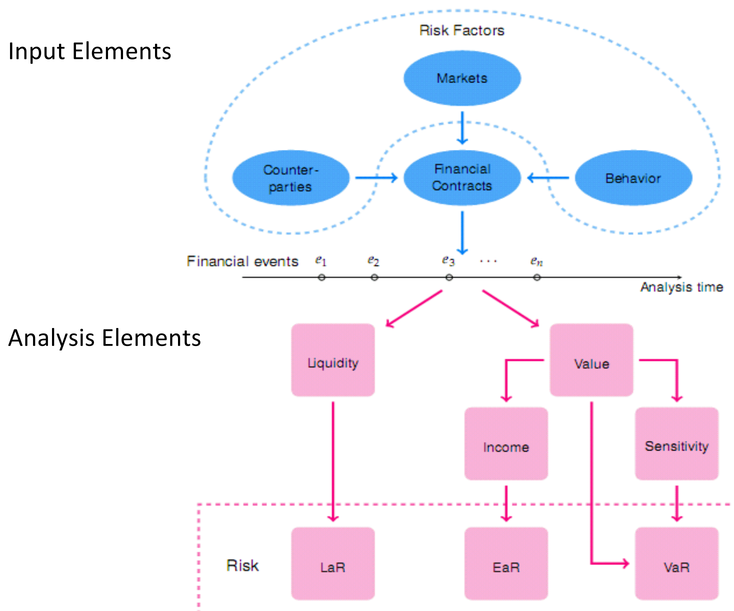


Figure 2: Base architecture

Figure 2 represents the architecture[4] needed to support the aims of the OFR, namely to "provide the data and the analytic tools to safeguard the U.S. financial system" or in the words of Senator Dodd "to look through the windshield" and not "through the rear-view mirror".[5]

The core of the model is the financial contract, a mutual promise of two or more counterparties to exchange cash flows according to a set of rules. The rules can be described by algorithms that determine their exact occurrence.[6] These promises form the only hard facts of finance.

[4] This text follows the article "The Regulatory Revolution* in GARP Risk Professional August 2010 by Brammertz/Mandelowitz. For more detailed arguments see also „Unified Financial Analysis – the missing links of finance" Wiley 2009, Brammertz et. al. especially chapters 1 and 2.
[5] GARP news 25.4.2010

Contracts are surrounded by and relate to the *risk factors* - markets, counterparties and behavior:

- Markets: Many contracts such as a floating rate bonds, options and so on, include rules which refer to market conditions such as interest rates, foreign exchange rates, or stock or commodity indices. Market conditions are also used for discounting and valuation.
- Counterparties: Although financial contracts represent hard facts regarding contractual commitments, the keeping of them is conditional. Counterparty data encompasses the conditions that determine the ability to meet the obligation.
- Behavior: Some rules governing the exchange of cash flows are not deterministic in the mechanical sense. The best known examples are saving accounts from which funds may be withdrawn at short notice or mortgages which in some countries can be prepaid without a fair-value penalty. Since such rules can only be formulated statistically they are part of the risk factors.

Market conditions, counterparty and behavioral information are called *risk factors* because only their current conditions are known, and they can change unexpectedly. Contracts and the risk factors are called *input factors* because they constitute the facts. The *financial events* can be calculated given any financial contract and their surrounding risk factors. In other words, the sequence of time-ordered *financial events* corresponds to reading the financial contract under certain risk factor conditions from which cash flows can be derived. In the next step, it is possible to derive the *analysis elements: liquidity*, *value*, *income*, *sensitivity* and *risk* (lower part of Figure 2). Note, that in such a system, *value* is not an input but an output. Value, in this system, can be calculated not only under current risk factor conditions but also under shocked or stressed conditions. Value can also be calculated according to any valuation principle (nominal, fair, amortized cost etc.), or in the case of options using different pricing models.

The distinction between *input* and *analysis elements* is crucially important. The classical manager is concerned with what we call here analysis elements. Liquidity, value and income constitute the three focal points of any financial manager. The classical manager is less concerned with the input elements, which are normally taken for granted.

However analysis elements depend fully on the input elements or – in other words – analysis elements are a handy way of expressing the state of financial contracts. However, once a value – be it book- or fair- or market-value – has been calculated the link to the causal factors is severed. Therefore: Input elements are important from an analytical perspective, while analysis elements are important from a managerial perspective and the link between them has to be maintained.[7]

It can be shown, that this system is complete from an analytical perspective inasmuch the financial contracts are represented in sufficient precision.[8]

This model greatly improves the regulatory data situation. While "old-style" regulators received in general pre-calculated outputs, mostly some kind of book or market value, the focus is now on the factors that determine those values. In the terminology of Figure 2, "old-style" regulation was focused on the analysis elements (lower pink part) while the OFR focuses on the input elements (upper blue part) and performs the analysis independently. With the upper blue past of figure 2 the regulator has the data necessary for robust analysis under various market

---

[6] It is common to think in legal terms in relationship to contracts which of course is also correct. For analytical purposes it is however more expedient to think in a set of logical algorithms.

[7] Suppose a market value of a certain bond is declared to be x. This is actually a short form of saying that a) the bond with a certain cash flow payment pattern under b) such and such market environment (interest rates, FX rates) and c) under the current expected probability of default has a value of x. Although the number x does contain all this information, it cannot – since compressed into a single number - reveal the underlying assumptions anymore.

[8] See Brammertz et. al. 2009 chapter 18.

and counter-party conditions. While in the lower pink part of Figure 2 the regulator has only a blurry snapshot in time under a fixed set of market and counter-party conditions from which no further analysis or insights are possible.

4. OFR AND DATA[9]

4.1. DATA COLLECTION

Of the four input elements, two are explicitly mentioned in the Dodd-Frank Bill: financial contracts and counterparty data. Market data (interest rates, FX rates etc.) are not explicitly mentioned, since these data are (rightly) assumed to exist in sufficient granular detail (Reuters, Bloomberg). Since the Behavior input element can only be described statistically (for example using prepayment tables), it is part of special research efforts and normally not directly collected from external sources. Thus it is not part of the data collection effort and not mentioned in the legal text, which however does not mean, that behavior is not important.

From the two data sources mentioned in the text and thus collected by the OFR from external sources, counterparty data – although tedious – is straightforward from the conceptual standpoint. In the core it is a unique identifier plus a relationship between the counterparties (i.e. a parent/subsidiary corporate relationship). In addition there is some descriptive data necessary to determine probability of default and correlation. The standardization of these data is just getting started with the current LEI (Legal Entity Identifier) project[10] and not further discussed here.

The crux is the contract data. As Figure 2 makes clear, contract data are hub of the system. Success or failure in this part of the mandate will determine success or failure of the OFR. We have to bear in mind, that the financial sector, operating on its own, has not yet found a satisfactory solution. Worse, banks struggle even internally with their multitudes of systems, each of them representing financial contracts in a non-standard way and spending billions of dollars every year in an effort to get consolidated overviews. The OFR can be imagined as a super IT department of all financial firms. How can it succeed in consolidating these multitudes of non-standard ways of handling contract data into an orderly overriding system?

4.2. STANDARDIZING FINANCIAL CONTRACTS

From its inception, it has been clear that the ability of the OFR to carry its mission rests on the standardization of financial contracts. Indeed, this standardization features prominently in the legislation, specifically requiring the OFR to build reference data basis in which all counterparties and contracts will be standardized for universal use.[11]

---

[9] Although presented in a conceptual style, the ideas put forward are not of a theoretical nature. They are based on twenty years of practical work and experience across almost all westerner countries and more than a dozen countries in the rest of the world. "Appendix 2: A simple data model and process" presents the ideas on a more practical level.

[10] There are surely sufficient challenges in fulfilling this responsibility such as complex relationships and ensuring that the data are kept up to date to involve a not insubstantial undertaking. However these challenges appear to be reasonably well understood and their solution straight-forward.

[11] See for example Sec. 153 (c), (2) STANDARDIZATION.—Member agencies, in consultation with the Office, shall implement regulations promulgated by the Office under paragraph (1) to standardize the types and formats of data reported and collected on behalf of the Council, as described in subsection (a)(2). If a member agency fails to implement such regulations prior to the expiration of the 3-year period following the date of publication of final regulations, the Office, in consultation with the Chairperson, may implement such regulations with respect to the financial entities under the jurisdiction of the member agency. This paragraph shall not supersede or interfere with the independent authority of a member agency under other law to collect data, in such format and manner, as the member agency requires.

However, standardization of financial contracts can take many forms and can have many meanings; therefore, we need to explain it further.

Consider the financial contract specified by the following data:

- Notional amount : 100 USD
- Value date : 1.1.00
- Maturity date : 31.12.04
- Interest payment cycle : 6 months
- Interest rate : 10%, 30/360

These contract elements may look familiar to anybody dealing with financial contracts. Many analysts may think that this contract is fully specified. Assuming the contract is an asset, most analyst would expect an outflow of 100 USD on the 1.1.0 followed by interest payment inflows of 5 USD each half year with a final payment of 100 USD at the end of the fourth year, leading to the cash flow sequence shown in Figure 3.

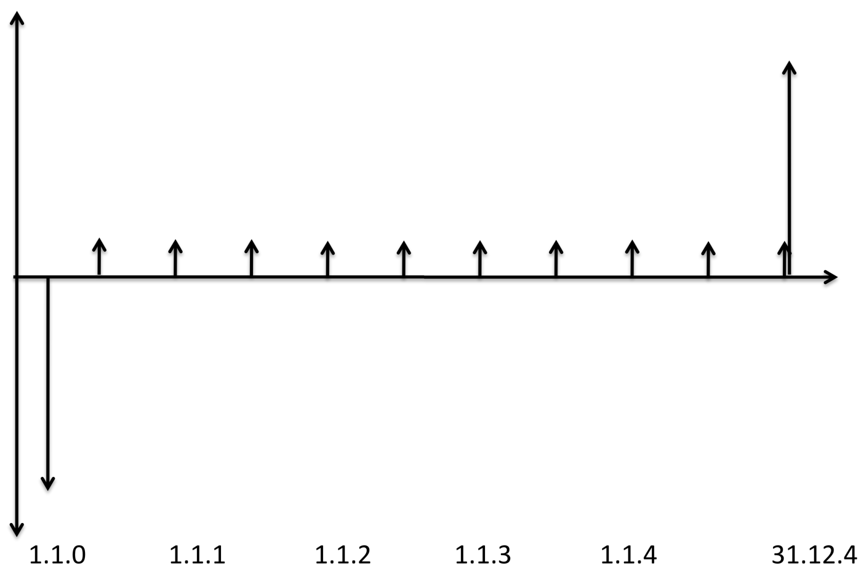


FIGURE 3: CASH FLOW SEQUENCE

From this expected cash flow sequence, the analyst could go on and derive value, say market value or any other book value of choice. Any other analytical concept from the lower part of Figure 2 could be derived and by combining them, any static ratio, risk ratio or risk-adjusted ratio found in any financial report could be calculated.

Although this representation sounds clear and logical, it could be totally wrong. Given the same data, an entirely different expected pattern of cash flows could be drawn as in Figure 4.

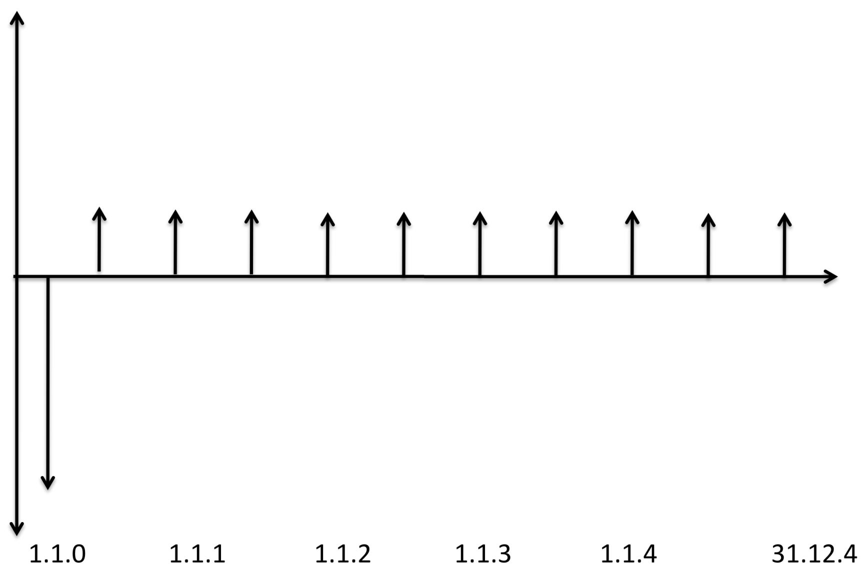


FIGURE 4: ALTERNATIVE CASH-FLOW SEQUENCE

In deriving the sequence of Figure 3, an implicit assumption was made about the cash flow pattern linked to the contract data: it was assumed that the data represent the cash flow pattern of a “bullet bond.”. However, given the same data an entirely different pattern is possible if the data describe an annuity. The resulting cash flow sequence is obviously quite different and as a consequence all analytical outputs (liquidity, value and so on) would be different. It is important to understand that the same set of data supports two different cash flow patterns. The only additional assumption needed for the annuity was that principal is paid at the same time as interest; a very common assumption made in many systems.

Other cash flow patterns based on the same set of data can be easily constructed. Instead of an annuity we could assume that principal is paid in regular tranches (for example 10 USD every half year), which leads to a decrease of interest payments from 5 to 4.5 and so on. Alternatively, the same data may represent a perpetual bond where the maturity date corresponds in fact to a call date. The implicit cash flow sequence in this case would again be very different. Such assumptions are not far-fetched. They happen in reality and are the reason why banks grapple with the problem and spend a great deal of money on it – with little success so far.

4.3. THE ROLE OF STANDARDIZED CONTRACT TYPES

We have shown that a unique interpretation of contract data cannot be achieved with better semantics at the attribute level alone since even when the semantics are fully specified it is still possible to derive different cash flow sequences. What is missing is a concept of contract types, or CT for short. A CT encapsulates the intended and legally binding exchange of cash between the two counter parties and turns attributes unambiguously into expected cash flows. Thus a CT consists not only of data attributes but also includes an algorithm that turns these attributes into expected cash flows. We could, for example, call the CT which generates the cash flow sequence of Figure 3 a principal at maturity or PAM for short, and the one which generates the cash flow sequence of Figure 4 an annuity or ANN. The two CTs of PAM and ANN stand for the attributes/algorithm combination that represents the intention of the contract in terms of cash flow patterns.

An objection could be raised at this point that the standardization problem is not really solved since the variety of financial contracts is so great as to approach infinity. Moreover, even if it was solved at a given point in time, financial firms are in the business of constantly inventing new instruments, which would quickly make any solution obsolete.

If this were really the case, our proposed solution would be indeed a pseudo solution. The good news is that, save for a small number of highly exotic contracts, the great majority of all financial contracts can be comfortably represented by a mere three dozen of CT's and with a high degree of precision.[12] Many of the new products are merely names, smoke and mirrors, cloaking a combination of simple products.

The question then arises why have the three dozen or so CT's not been defined a long time ago, in particular since it seems so easy? Was it inability or unwillingness, stupidity or wrongdoing? Many would argue that unwillingness is the root cause since opaqueness in financial products is profitable for sellers. Indeed, bankers were accused during the 2008 crisis of promoting opaque products for their own benefit.

The argument certainly holds some truth. Despite this, we think that inability played an even more prominent role. Since the advent of the computer in the financial industry, banks have been haunted by internal and external (mainly regulatory) demands, forcing them into a sequence of makeshift solutions and devoting too little time to ponder fundamental architectural questions. The ever shorter reporting cycles made it even worse.

Then why should this change now? The answer to this question is the OFR itself which is the first institution mandated to tackle this problem. The OFR has the power to set standards that go beyond mere semantics. Once these standards are set, they will not only benefit the OFR and other regulators but will ultimately benefit the financial industry. It will be a similar achievement as when DOS was introduced by IBM: it created a de facto standard in a previously fragmented industry, which yielded benefits for buyers and sellers.[13]

5. THE CORE PROBLEM

If the notion of CT's is so important, then why not just add one more attribute to each financial contract to any database and solve the problem? Although this would be a step into the right direction, it does not address the core of the problem.

In order to follow this argument, we have to start with the separation of data and algorithms in the natural sciences, followed by a discussion of how this separation should be adapted to the financial industry context. Then we must discuss where this separation makes sense and where it should not apply. Our chief conclusion is that the separation is disastrous for the concept of CT's but is needed for the modeling of risk factors.

5.1. SEPARATION OF DATA AND ALGORITHMS IN NATURAL SCIENCES

That data and algorithms should be separated is commonly accepted as sound principle.[14] For example the "Data Processing User Guide" of the European Laboratory for Particle Physics states it as follows:

[12] See Appendix 1: exotic products for a discussion of the outliers. In short it can be said, that they do not pose a serious problem.

[13] It has been pointed out, that this is a poor analogy since DOS is considered a bad standard by many and that the OFR should not follow this example. While I agree with this objection, I have to stay with the example lacking a better one. What was really achieved, even though with a bad standard, was a defragmentation of a fragmented market.

[14] Gaudi User Guide p. 8, European Laboratory for Particle Physics, CH-1211 Genève 23 – Suisse.

https://lhcb-comp.web.cern.ch/lhcb-comp/Frameworks/Gaudi/Gaudi_v9/GUG/GUG.pdf

"It is intended that almost all software written by physicists, whether for event generation, reconstruction or analysis, will be in the form of specializations of a few specific components. Here, specialization means taking a standard component and adding to its functionality while keeping the interface the same. Within the application framework this is done by deriving new classes from one of the base classes:

- DataObject
- Algorithm
- Converter"

The main separation is between data and algorithms; the third category is of minor importance.

The ostensible reason for this separation seems to be robustness and simplicity. However, there is a more profound reason behind this choice of architecture in natural sciences: physics and all natural sciences are in a continual experimental state. Actually science is interesting only in fields, where new conclusions can still be expected. This demands continually new hypotheses applied on more or less the same data. Mixing data and algorithms under these conditions is obviously not a good idea. Keeping on one hand the facts (the data), which can be considered stable and on the other hand the possible conclusions to be drawn (hypotheses), which are not necessarily stable, makes much sense.

5.2. SEPARATION OF DATA AND ALGORITHMS IN THE FINANCIAL SECTOR

The paradigm of separation of data and algorithms seems to have slipped into the financial circles as an indisputable truth without need of much reflection. Most, if not all, banks follow this model. It became especially enshrined within and by the Data Warehouse movement. The separation is considered smart also for semantic repositories since it allows for "multiple levels of abstraction, from multiple sources and with multiple views"[15]. As a consequence, such databases care about data but not about algorithms, which are considered subjective and left to the eye of the beholder. An example often quoted by specialists, why data and algorithms have to be kept separate, is valuation. There are so many ways to calculate value, who then can proclaim the single truth?

Is this separation within the realm of finance as clever as it is proclaimed? The answer: yes there are parts, which are highly subjective within finance. In these areas the separation of data and algorithm is of great help. However equally true is the fact that there are many parts – possibly the larger parts – in finance, which can be considered mechanical with only one single interpretation. Now we have the worst possible situation: both parts are mingled to the point where they have become indistinguishable. The two spheres having become indistinguishable produced the effect that even the mechanical parts are treated often as subjective and separation of data and algorithms is applied throughout.

We will discuss now which parts are mechanistic and need to be treated as such, which parts are experimental and finally how they work together.

5.3. THE "MECHANICAL" PARTS OF FINANCE

The intention of a financial contract – be it a mortgage, a loan, a swap or an option – is to clearly define the rules which govern how cash is exchanged between two parties. The rules might be simple, as in the case of a fixed bond or complex like in a ladder option. Nevertheless, the intention is always to leave no room for interpretation. If the contracts were not defined clearly enough, it will be left for lawyers to close the gap. Here we discuss the case

[15] Designing a Semantic Repository, Integrating architectures for reuse and integration by Cory Casanave, May 2007 (http://www.w3.org/2007/06/eGov-dc/papers/SemanticRepository.pdf)

wherein the interpretation is clear, which is the bulk of the contract calculation. In the next section, we will discuss the exceptions.

Let us return to Figure 2. In the center of the blue (upper) part of the diagram we have a contract surrounded by the risk factors, which have at any point in time a certain state, a state that includes expectations such as those expressed in volatilities. A clearly defined contract connects these risk factors uniquely with the expected financial events and finally the expected cash flows. Given the expected cash flows and a bookkeeping rule – which can be mark to market or any other – all the analytical elements are clear, except the risk part.

Since the relationship between the financial contracts and a given state of the risk factors yields a unique sequence of cash flows, it is possible to represent it by a unique code. The code represents nothing other than the intentions of the creditor and debtor, or more broadly the two parties to the contract. Since it is possible to reproduce the cash flow pattern of an overwhelming majority of existing financial contracts – including most exotic products – by a mere three dozen CT's, defining these CT's once and for all would provide a significant efficiency boost across the financial industry. If CT's were well defined and understood by all market participants, any analyst could uniquely derive the expected cash flows. From the unambiguous cash flows, other variables of interest, such as the value and income of any portfolio of contracts can be derived as a function of the current state of the risk factors.

The quantum leap achieved with this approach can best be understood if we contrast it with the current approach applied in most banks today. Based on a strict separation between data and algorithms, we find more or less well-described input data within databases. It has to be noted, that this data coincides with the upper part of Figure 2. However, what is done with the data is left to the beholder. This is the case irrespective of the fact that any financial analysis depends on the expected cash flows and that the contracts are defined in such a way as to return a unique sequence of cash flows, given a particular state of the risk factors. This also does mean that even though different analyst may ask different questions, they really intend to model the same underlying expected cash flows, given the same state of risk factors.

In practice, it is only by coincidence that the expected cash flow streams calculated separately will turn out to be the same. In general they will diverge due to cost pressures, which inevitably lead to shortcuts such as making different approximations. Another reason for the divergence is that absence of strictly defined CTs. This leaves too many open questions and one programmer might derive the cash flow sequence as in Figure 3 and another as Figure 4, although only one sequence can be the intended one.

Every new analytical question tends to trigger a new project within banks. New projects end up with new pieces of code generating implied cash flows, a futile and endlessly repeated exercise. Based on hundreds of projects in banks, my personal estimate is that 70 to 80 percent of such work could be avoided. Moreover, the quality of the analytical output will improve since it is worthwhile to invest once and for all in a well-defined engine that returns reliable expected cash flows.

Only the abolishment of the strict separation of data and algorithm will solve the problem. The concept of CT's defines not only the data but incorporates at the same time the exact algorithms for deriving the cash flows given a state of the risk factors. This will put an end to the endless reinvention of the wheel from which not only the OFR, but the industry as a whole will profit.

5.4. THE SUBJECTIVE PARTS OF FINANCE

In the previous section we made a strong assumption about the risk factors in as much that everybody agreed about implicit expectations. Everything about risk factors was assumed to be clear, even their implied evolution.

This assumption is certainly valid for forward market rates. It is less valid for volatility term structures and certainly only partially valid for correlations. Its validity is less than universal for credit ratings and migration matrices and even more so for the behavioral elements. It is not valid if risk is the analytical question. In short, risk factors are risky.

Disagreement about valuation, especially if it comes to fair or market values are essentially disagreements about the evolution of the risk factors. Assume a different path and we will get a different valuation. Different option pricing formulas for example are different ways and assumptions about the future evolution of the risk factors. However: if two parties agree about the evolution of the risk factors then the result must be the same. Going back to our examples of Figure 3 and Figure 4: only one picture represents the intention of the two parties.

Does this devalue what has been stated in the previous sub-section? No. Even if people disagree on the future evolution of the markets, a well-defined contract still gives unique results given any assumed evolution. We might disagree on volatilities and correlations however we cannot disagree on the expected cash flows once volatilities and correlations are chosen for a well defined contract. This part, which is still the bulk of the complexity, remains mechanical and thus should be solved in a mechanical fashion without any unnecessary degrees of freedom.

As a corollary, this implies however, that separation of data and algorithms makes sense if it comes to risk factors. It must be possible to define the most diverse models of risk factor evolution. Such models might be highly scientific or ad hoc like in stress tests. Once set however, the mechanical part of cash flow generation kicks in again.

Besides risk factors, there is a second class of uncertainty to be dealt with: Financial contracts that are not well defined in terms of pay off. Within this category, two different categories are to be distinguished: First the problem of “fuzzy” contracts such as savings and mortgages which allow prepayment (in some cases) and second clearly defined contracts where the rules get broken.

Classical saving accounts are an example for fuzzy contracts, where neither a clear maturity date nor a rate reset pattern is set. Maturity is up to the holder of the account while rate adjustments are up to the bank to decide. They often follow market conditions but in a very imperfect fashion. Prepayment of mortgages as practiced in the United States is another example. In both cases we do not have a precise rule that defines expected cash flows, which complicates financial analytics vastly.

Getting around this problem demands assumptions about the behavior of the counter parties to these contracts. For example how savings account holders react to interest rate differentials between the savings rate and alternative products or how they react to market rumors concerning the bank offering savings products. Prepayment models predict prepayment speed as a function of rate differentials, age, remaining term of the loan and so on. Such models are based on some statistical observation and are far less stringent than for example term structure models. They change often depending on new data and insights.

Breach of a contract is the other source of uncertainty related to contract rules. Credit risk is a common case and is handled via rating, migration matrices and probability of default as a part of the counterparty data. More complex are cases where contract terms are modified by unforeseen external circumstances. This happened for example in the wake of the sub-prime crisis when the US government “encouraged” banks to extend the teaser rates on mortgages for an additional year or two. The original contract was clearly defined, however the government intervened to create an unexpected change.

Both cases are solvable via a "behavior" component in the analysis where it is possible to write functions that directly address the conditions of the contracts. An example would be: "Take all contracts of the type ANN which belong to a certain counterparty class and whose value date is less than two years in the past and postpone the next rate reset date by one year". The execution of the function would affect all cash flows of all contracts consistently and the effect on value and income could be calculated immediately.

5.5. COMBINING THE MECHANICAL AND THE SUBJECTIVE

The interaction of the two parts can again be conveniently explained along Figure 2. If risk factors are considered as known (or at least their expected values), then the rules encoded in the CTs transform contracts into sequences of expected cash flows from which further rules turn them unequivocally into the liquidity, value, income and sensitivity analysis elements. In this sense the entire system can be considered as mechanical. In such a world it is sufficient to apply the current state of the risk factors to the financial contracts. Since each contract is mapped to a specific CT, the algorithms embedded within the CT will perform the tedious work of turning the data into expected cash flows that further algorithms turn into the analysis elements.

Although practitioners do work routinely under efficient market assumptions, it has been shown that expectations can change rapidly and in unexpected fashion. Therefore a necessary condition for such a system to be successful is to keep it "experimental" in terms of the risk factors. It must be possible to apply different algorithms modeling the evolution of the risk factors and especially, it must be possible to form subjective stress scenarios.

Once these expectations are formed, the system turns again "mechanical" in operation. Expected cash flows are generated and the analysis elements derived. New values, incomes etc. are derived under the new assumptions. The cycle can be repeated as often as needed. What we win with this approach is a risk management system worthy of its name and applicable from a single financial contract up to the systemic level.

Thanks to the limited variations of the bulk of the financial contracts, this high quality risk management system can be implemented for an affordable price. Since the bulk of the system is mechanical, it can be programmed once with finite effort. Instead of grappling with data, the analysts can focus on grappling with the problems they should be dealing with: with the risk factors and various assumptions that must be drawn and tested.

6. CONCLUSION

The power of the presented model lies in its clear structure, which separates the mechanical from the experimental parts of finance. The mechanical parts are encapsulated within the concept of CT's, which combine attributes and cash flow generation in a unique and well-designed manner. The experimental part is encapsulated within the risk factors and allows any modeling to represent different models, assumptions or even subjective expectations. The relationship between risk factors and contracts is again mechanical.

The model also starts clearly from the input elements and proceeds to analytical outputs.[16] This empowers the OFR and therefore regulators, and finally the government as the party charged with oversight of threats to financial

[16] Although this should be the only obvious way of dealing with such problems, it is in no way a common practice in financial institutions and especially not in regulation. Traditional regulation does not start with the input elements and then progress towards the analysis elements. It rather demands analysis elements (mostly a book- or a market-value) from the member banks in order to sum them up consequently. OFR demands a shift from the analysis elements to the input elements in order to empower the regulators to perform any analysis themselves.

stability, to perform independent analysis of any portfolio, any bank and even the entire financial sector on a system-wide level. Analysis includes all known concepts especially any risk analysis and stress scenarios. This is the look through “the windshield” promised by Senator Dodd.

Not only the OFR, regulators and the government will benefit from such a system but the financial industry itself. Financial analysis will become much easier and will cost drastically less. This will be partially offset by expected losses due to demystification.

The concept of CT’s is of crucial importance. CTs link the contractual data uniquely with the intended expected cash flows. Without such a concept the OFR will be inundated by undecipherable data in short order even if each single attribute is well described.

Not only are regulators liable to this type of reversed engineered thinking. Accountants alike start their thinking process with the balance sheet that represents “value”. Even modern finance falls into the same trap: courses on modern finance tend to start with a professor writing something like the Black-Sholes formula starting with C(S,t)= …., the value of a call option.

7. APPENDIX 1: EXOTIC PRODUCTS AND STANDARDIZATION

Although it is possible to model the vast majority of financial contracts with a relatively small set of standardized CT's, there will always be outliers. This is not only true for the methodology described above here, but holds for any analytical methodology. The very notion of standards implies the existence of outliers or, put differently, where there are standards, there are also exceptions.

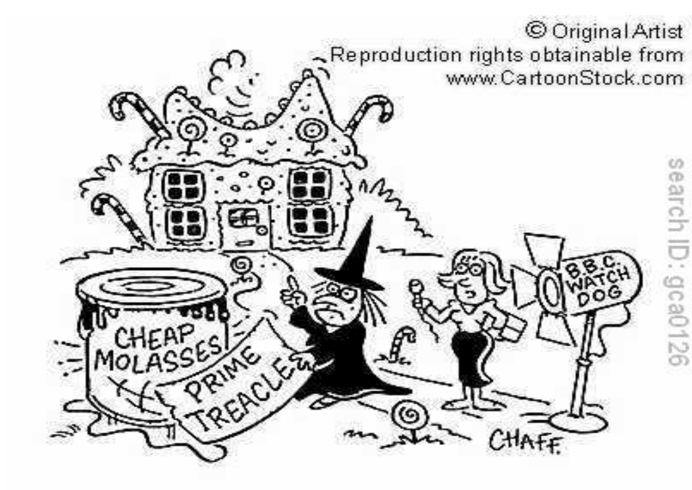
In this appendix we discuss the state of standardization in the financial sector and the technical solutions to the problem.

7.1. STANDARDIZATION IN THE FINANCIAL SECTOR

It is common to define standards in public sector areas. Air, water, electricity, transport and food are examples of important public sectors, which are more or less regulated. After all, these are public goods we all depend on. If I drink a glass of water I want to be sure, that certain standards be adhered to.

Standards, if they are to be adhered to, require language through which they can be communicated. In the food industry for example it is the language of chemistry, together with labeling standards for packaging. The declaration and the contents must match within narrow margins of tolerance.

Although the financial sector can be considered to be one of the most important public sectors and although it is much more difficult to describe food than the conditions of financial contracts, there is nothing comparable to a food declaration in the financial sector today. The imperative of food declaration - declaration and content must match – does not apply in finance. Worse even, the financial sector simply lacks a comparable language to even describe the content of its products. The declared intention of the OFR is to change this. To do so, it needs a concept of CT's.



The financial sector is in terms of standardization where the industrial sector was in the middle of the 19th century, when each screw had its own nut. Despite this desperate state, there is a tendency in the financial sector to undo any standardization effort by pointing to the outliers. Much of this criticism seems to be driven from a protective mode. Whatever the reason behind this tendency[17], we should remember that:

- The vast majority of the contracts do indeed follow standards even though undeclared and thus foregoing its potential. Standards are a hidden fact.
- By obstinately focusing on the notorious difficult to tame corners of finance, we forego solving the vast majority of the cases, where this is easily possible. By not be able to fix the 1%, we give up on the 99% which would be easily fixable.
- The SFTC and SEC prove that standardization is possible even in the "notorious corners" of finance. They do so by forcing OTC derivatives through a clearing house. OTC derivatives such as CDO's and complex options have hitherto been considered the most buoyant and innovative part of finance and the most difficult to control. If it is possible to bridle this part, there is hardly an argument left against bridling the rest.
- Even if it would be impossible to ever find a solution for the outliers: A system which is capable of encompassing the large majority of the contracts would still be a major accomplishment for the OFR, even if the rest are handled by an approximation, which always is the case. Assuming that we can model 99% of the financial contracts to a high degree of precision with standard contracts and the remaining 1% with an approximation of 70%, we still would get highly precise overall results. This is even more the case when considering he heavy assumptions that need to be made in the area of the risk factors.
- Comparing the manufacturing sector with finance: Would anybody like to go back to the state, where each bolt had its own nut? What was the value of each nut having its own bolt? Nothing. It is very likely that a few years hence we will view non-standardized financial products in a similar light as we view individual nuts and bolts today.
- Finally: It is even possible to standardize non-standard product from an analytical perspective. This we discuss next.

7.2. Standardization of the outliers

Although sounding paradoxical, it is possible to standardize even outliers. From Figure 2, we can deduce that if it is possible to generate the financial events it is possible to derive the analysis elements. Generation of financial events in an unrestricted world however is difficult or impossible wherefore we have to work with cash flows directly. By skipping the event line which contains the full information for the liquidity aspects and the value/income/sensitivity aspects of finance, we have to salvage the system by producing directly the expected cash flows in two forms: one for the liquidity side and one for value/income/sensitivity[18].

[17] From my personal experience I found generally two intentions behind this tendency:

- Opposition to Standardization: Standardization is the enemy of opaqueness. Since opaqueness pays in the financial sector, standardization will be opposed by some who profit from opaqueness.
- Intellectual curiosity: Quants like exotic products. I was in many sales situations where the topic quickly turned to exotic products and how the system can handle it. Most of the time was consumed by this topic. When asked how many of the products discussed they had on the books the answer was invariably either "a very few" or even "none". Besides intellectual curiosity such discussions were driven by the assumption, that a system that can handle exotic options can handle simple products, as a matter of course, which is a fallacy. I haven't seen a system specialized in exotic options that can handle the day-to-day loans or even saving products.

[18] The following argument is taken from Brammertz et. al. 2009 page 69f including the figure.

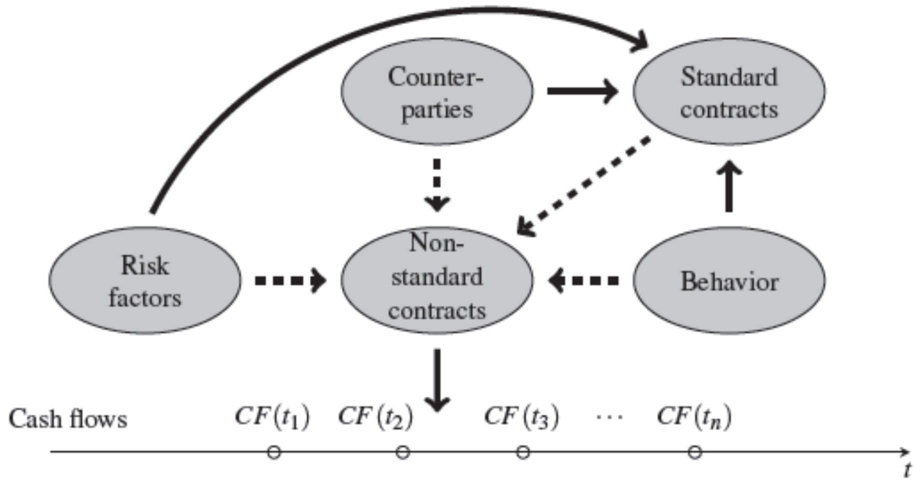


FIGURE 5: STANDARDIZING NON-STANDARD CONTRACTS

The system is like a mini-version of the whole system. Figure 5 is almost the same picture as the upper part of Figure 2. There are however four notable differences.

- Nonstandard contracts instead of financial contracts: This implies using a general purpose programming language to define the cash flow rules as necessary.
- Dashed arrows: The dashed arrows between the nonstandard contract and the other three input elements stand for the functional relationship between these elements. The output of the contract necessarily depends on the other input elements, including in particular standard contract types. If we think about an exotic contract any relationship is feasible such as dependence on the difference of the three-month interest rate and the mean temperature in Florida, or the difference between the market value of a certain bond and a stock. Ratings, prepayments (as part of behavior) and many other factors can play roles, which are represented by the input factors. A nonstandard contract may depend even on other standard (or even nonstandard) contracts for example in the case where the payoff is a function of the value difference between a stock and a bond. Such standard contracts depend of course in a consistent fashion on the input elements. In order not to overload the picture the arrows indicating this dependence are left out.
- Output is pure cash flow: While standard contracts produce about two dozens of different event types, there is only one event type known to the nonstandard contract type: cash flow as such. In order to cover the full analytical spectrum it is necessary to write two pieces of code: one generating cash flows for liquidity related analysis and one for value/income/sensitivity related analysis.
- The cash flow generating code has to be delivered along with the product.

What is standard in this system is the structure of the process, which links risk factors, financial contracts and cash flows.

Since the code has to be written for each product, which is an error-prone process, additional care has to be taken from a regulatory perspective. One precaution would be additional checks and another, additional capital charges.

Finally: Non-standard products can easily change into standard products and thus reducing the additional capital charges.

7.3. THE BOUNDARY BETWEEN STANDARD AND NON-STANDARD CT'S

The set of the possible contracts is not fixed or immutable, but must expand over time if new relevant instruments appear on the market. Where standardized contracts end and non-standardized contracts begin, is a question of choice. The OFR could set itself for example a target to model at least 99% of all relevant contracts with 99% precision – where relevance could be measured by a value, sensitivity or risk measure and precision by correctness of amount and timing of the expected cash flows. This would demand a study of all existing market instruments. We estimate, that 99% of the current existing variety of contracts can be covered with a mere three dozen mentioned CT's on a 99% precision level.

In 8.5 we will discuss shortly the additional effort non-standard contracts do take. We will conclude there, that with a system proposed as this one, the likelihood of appearance of new non-standard is reduced and where they still will appear, they will quickly mutate into standard products.

8. APPENDIX 2: A SIMPLE DATA MODEL AND PROCESS

8.1. INTRODUCTION

Data models often exhibit the following characteristics:

- Data models tend to become overloaded with hundreds of tables, attributes and unclear relationships.
- Financial databases are often organized along trading and legal considerations, where the product category (e.g. derivatives, equity, loans, debt and so on) is often the primary discriminating principle.
- The separation along legal and traded/non-traded lines has the effect, that the same attributes are defined in several tables often with slight or even more pronounced unintended definitional differences, adding yet another layer of complexity and difficulty.
- Too strong emphasis is placed on tradable instruments at the expense of on-balance sheet instruments, which are often aggregated into a single group despite the huge variance found in this segment.[19]

Fortunately, the situation can be improved if data are organized around the idea of cash flow generation. In the following we will sketch a simple entity relationship model with the financial contract at its center. This will be followed by a discussion of the cash flow engines linked to the contract types and some practical issues.

8.2. ENTITY RELATIONSHIP

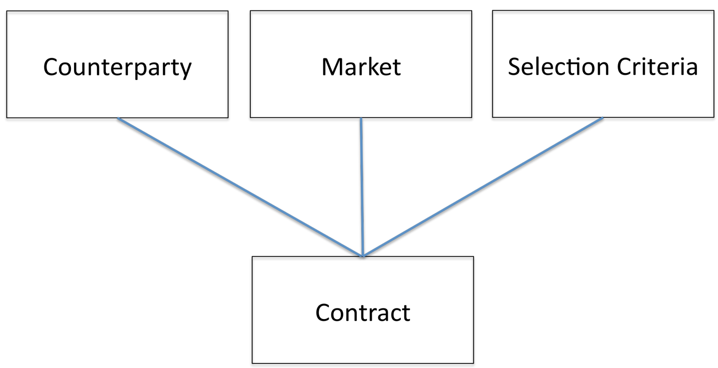


FIGURE 6: ENTITY RELATIONSHIP

Figure 6 shows the entity relationship of the model in its simplest form. At its center is the (financial) Contract, which is linked to the Counterparties, Markets and Selection Criteria entities. This figure is similar to Figure 2 save for Behavior and Selection Criteria. There is no relationship between Behavior and Contract due to the statistical nature of Behavior, which does not allow treating it as simple exogenous data.[20] Behavior is rather represented by algorithms with a constant shifting set of parameters. Selection Criteria – for example profit centers, product names, business units and so on – are on the other hand factual data that does not directly influence event / cash flow generation but can be important in analysis and reporting where they are used as grouping criteria.

Since market and counterparty data are not our focus, we keep the representation as simple as shown here and do not further expand it. Is it however possible to represent financial contracts by a simple table? My experience has

---

[19] This is probably an effect of the strong presence of trading oriented personnel in such exercises.

[20] Most Behavior models – for example withdrawal patterns of saving accounts – are proprietary due to the individual customer base of each bank. Nonetheless there are exceptions, especially in the area of prepayment of long-term fixed-rate U.S. mortgages.

shown that a single table indeed best represents most financial contracts. There would be possibilities to fragment a single table into sub-tables such as for principal payment attributes, interest payment attributes and so on. However practically speaking, such tables increased readability only slightly at the best and at worst, they decreased computing performance. There are however some special cases in the area of exotic options with very rarely used attributes, where sub-tables can be considered.

A basic contract such as a loan, a discount paper, a deposit or a stock is simply a single record in the Contract table. Combined contracts such as swaps, futures and options are represented by two or more records, where one record corresponds to the parent contract (e.g. the option) and the additional records correspond to the underlying(s). An interest rate swap for example will be stored using three Contract records, with the first representing the swap agreement as such and the two other records corresponding to the two legs of the swap.

8.3. THE CONTRACT TABLE AND VALIDATION

The contract table has about 150 financial event / cash flow determining attributes. The 150 attributes cover all cases of all the CT's. In it's simplest form the table can be imagined to have 150 columns or attributes along the top of Figure 7 with each row representing a contract or a sub-contract. Combined contracts such as swap would occupy several lines.

It has to be noted that in this table each attribute appears exactly once. This approach avoids the multiple parallel presentations of the same or similar attributes found in many databases and guarantees a single and precise interpretation. For example, the principal of a bond or a loan or the underlying of an option are defined and interpreted exactly the same way for every contract.

Record number	Contract Type	Principal	Principal payment cycle	Principal payment start	Interest rate	Value date	Maturity date	Repricing start	…
1	PAM	1'000	-	.	3.5%	1.5.03	1.5.13	-	
2	IRSWP	-	-	-	-	20.9.08	20.9.14	-	
2a	PAM	-500	-	-	2.8%	-	-	-	
2b	PAM	+500	-	.	0.95%	-	-	15.7.06	
3	ANN	-2'000	1M	1.1.07	4.1%	1.10.10	NULL	1.10.20	
…									

FIGURE 7: CONTRACT TABLE

Figure 7 shows a tiny section of the fields. Each column represents an attribute and each row a contract. The most important attribute is the CT, which finally defines the cash-flow generation. The first row is an example of a principle at maturity, the second of an interest rate swap (IRSWP) with its two legs on row 2a and 2b. Row 3 is an annuity and so on.

While some of these 150 attributes are commonly shared between many CT's, most of them are not shared by all. Some CT's use them, others not. Actually many CT's can be described with a few dozens of them. Figure 8 shows a table describing this relationship where we have on the row headings the same attributes as in Figure 7 and on the column headings the CT's.

	PAM	ANN	...
Principal	NOT NULL	NOT NULL	...
Principal payment cycle	NA	NOT NULL	...
Principal payment start	NA	NOT NULL	...
Interest rate	NOT NULL	NOT NULL	...
Value date	NOT NULL	NULL	...
Maturity date	NOT NULL	NULL	...
Repricing start	NA	NA	...
...	...	...	...

FIGURE 8: CONTRACT VALIDATION RULES

The cells of the table describe whether an attribute is compulsory (NOT NULL), whether it is not applicable (NA) or whether it is applicable but can be left empty (NULL). Nullity can be conditional, depending on the existence and value of certain other fields. Theses are validation rules. Additional validation rules apply including consistency criteria. Validation rules are CT sensitive.

Independent of where and which database the data would come from, Figure 8 would define the data standard. The chosen format is of secondary importance, however the data exchange with the OFR must allow for semantic as well syntactic validation of the submitted data. This could be done for example using XBRL.

Most important in this exercise would be a clear and concise communication with the financial sector. The following minimal communication conditions have to be met:

- A complete set of the CT's as defined in 7.3
- A complete set of attributes
- A table along the lines of Figure 8
- Clear semantics, including
    - A clear and understandable description of each attribute from a financial perspective
    - The technical format of each attribute
    - Nullity conditions per CT
    - Interpretation of Null per CT
    - Consistency rules per CT

Although a good start, this is not the end. As we argued in 4.2, precise semantics on the attribute level alone is not sufficient. We need clear semantics on the level of event and cash flow generation.

8.4. THE CASH FLOW ENGINE

The data described in the previous section represent the input data in the upper section of Figure 2. Although indispensible for the type of analysis, which the OFR should perform, it is neither the ultimate aim nor the end. From an analytical perspective the lower –analytical – part of Figure 2 is of ultimate interest. To reach there, one must pass through the event / cash flow layer in between inputs and analytical outputs. The engine that connects input data and final analytical results is often referred as the cash flow engine.[21]

A reference implementation of an event generation engine should be made available. The following minimal standard has to be met:

- Publish examples of real life contracts (with their diverse names) and show they map into the standard CT's
- Document
    - Cash flow generation per CT (along the lines of the examples in Figure 3 and Figure 4)
    - Sub-mechanisms like interest calculation, principal payment, option settlement and so on
- Make available a reference implementation in source and binary forms with permissive license terms that allow unencumbered use.
- Make available an interface where contract data can be submitted and the corresponding cash flows produced by the reference cash flow engine are returned

Once the financial events and cash flows are determined, there is a final step of deriving the analysis elements liquidity, value, income and sensitivity and – where applicable – risk.[22] These steps can be formulated independent of the CT's[23], should however be part of the total package.

In summary: we need a well-published cash flow engine plus an engine that turns financial events into liquidity, value, income and sensitivity. Everything has to be published and made publicly available via easy to understand concepts and support tools. The system is complete.

8.5. THE MAPPING PROCESS

Assuming information is available on the attribute and contract level as just laid out, the participants can start their mapping process. Mapping means turning the data from an internal representation into the target standard representation. How the data is organized is of minor importance as long as it can be retrieved. Such a mapping process has the following main steps:

- Decide how each financial contract maps into a standard CT. This will be in most cases a one-to-one representation if the set of CT's is well chosen. In some cases however it would be a one-to-many relationship, where one financial contract must be represented by more than one standard CT. In few cases even a many-to-one relationship may be needed.

[21] Event generating engine would be more appropriate, however we stick with the commonly used term cash flow generator. Events – in contrast to cash flow – contain the combined information covering liquidity, income and value while cash flows can only represent either liquidity or income/value.

[22] Risk calculation means in many cases moving or shocking the risk factors and recalculating liquidity, value, income under shocked or Monte Carlo conditions. In some cases, such as parametric VaR it can be directly deducted and published as a code.

[23] This is obvious for liquidity, which is pure cash flow and independent of any valuation or bookkeeping rule. Value and income related concepts need in addition a bookkeeping rule. Once the rule is set – and there are not many bookkeeping rules in the financial world – the result is again unanimous.

A contract is well mapped, if the minimal requirements of the OFR are met. As discussed in 7.3, these requirements could specify for example a certain precision in terms of size and timing of the expected cash flows.

- Once the choice of the CT's is clear, the standardized and publicly available (complete) list of Figure 8 specifies which attributes have to be mapped per CT.
- The data are provided to the OFR in a standardized format.

The OFR will be in a position to interpret the data correctly and to test hypotheses.

A short word yet on mapping of non-standard contracts. In order to communicate non-standard contracts, the same process as for the standards has to be gone through, however on individual and repetitive basis. If there are non-standard attributes, they must be clearly described and communicated with the potential consumers of the data. A cash flow generating code must be made equally available. This is obviously a tedious job and one of the following results is to be expected:

- There will be less non-standard products, since their additional value is often small or nil and does not justify the additional efforts. The need to openly declare it as a non-standard will have further deterring effects on the demand side. Special regulatory charges for such products could discourage the use even further.
- If the non-standard product is still considered valid, then it is likely that the industry will add this CT as a standard, thus keeping the system current as the market evolves.