

Interaction-driven Definition of e-Business Processes

Christian Zirpins

*Dept. of Computer Science (VSYS)
University of Hamburg
Hamburg, Germany
zirpins@informatik.uni-hamburg.de*

Giacomo Piccinelli

*Hewlett-Packard Laboratories
Stoke Gifford Park
Bristol, BS34 8QZ (UK)
giacomo_piccinelli@HP.com*

1. From Workflows to e-Business Processes

Business-to-business interaction (B2Bi) is the next step for corporate IT [1]. Business relationships become increasingly dynamic, and new requirements emerge for data and process management. Standardisation initiatives are successfully targeting business ontology [4]. Still, business agility mainly depends on the flexibility of the business processes of a company. In the B2B space, traditional approaches to process modelling and management are inadequate.

Today more than ever, traditional workflow management is crucial for the internal effectiveness of a company. Internal efficiency is a prerequisite for external agility. From both a technical and a business perspective, internal workflow management relies on specific assumptions in terms of resources involved in the process, as well as the process itself [2]. Level of control, availability, reliability, and cost stability are parameters that traditional process models and technology can almost take for granted. A single authority ruling on the process definition and the total control over process execution are also basic concepts for internal workflows. From a business perspective, a big upfront investment is put in the complete definition of process specifications. A different conceptual framework is required for the definition and management of e-business processes [3, 5]. The intrinsic capability to adapt to rapidly changing business requirements becomes crucial.

The line of research explored in this paper derives from an approach to process modelling and management that explicitly targets the peculiarities and dynamics of B2Bi. In the model we propose, the upfront specification of the interaction logic of a company can be limited to partially specified processes and basic interaction rules. Specific information is then gathered from the observation of actual instances of business interaction, and used to refine and extend the initial model. In addition to the enforcement of explicit business requirement, the goal is to capture and leverage implicit operational knowledge.

In the following sections, we present an overview of the methodology we are currently experimenting with for the inference of complex processes from business interaction flows. For our initial experiments, we focus on

business messages compliant with the RosettaNet standard [4].

2. Dynamic discovery of interaction logic

Business interaction represents the external manifestation of internal business processes. On the one hand, interaction logic is regulated by specific business rules. The body of rules governing the information exchange shapes the interaction in accordance with specific patterns. On the other hand, different parts of a company use different processes. Different rules are applied in different contexts. For example, the notification process used by the finance department can be quite different from the one used by the sales department. Formal specifications and enforcement mechanisms can be quite different, if at all available.

In an ideal-case scenario, all the business processes of a company are captured in formal specifications, and enforced by some form of workflow management system. Initiatives such as the ISO 9000 certification programme are an attempt to move businesses in such direction. In practice, cost-effectiveness evaluations often discourage companies from formalisation and automation. Workflow management technology can be quite expensive, and the cost of the technology is just a fraction of the costs of process engineering. Individual applications automate specific aspects of different business processes, and ad-hoc solutions are usually put in place to achieve coordination. The actual logic behind business interaction emerges from diversified reality of applications and procedures. Direct communication is one of the few points of convergence for interaction logic.

In practical terms, different data-mining techniques and information models can be applied to the communication channels of a company in order to derive information on interaction patterns. For the purpose of our work, we identified two major requirements. The first requirement is that the information model has to focus on individual business transactions. Given the messages related to a transaction, the model should represent the type of causal and temporal relations used in static business-process definitions. The second requirement is that the interaction process associated to a specific transaction has to be derived at the same time in which the transaction unfolds.

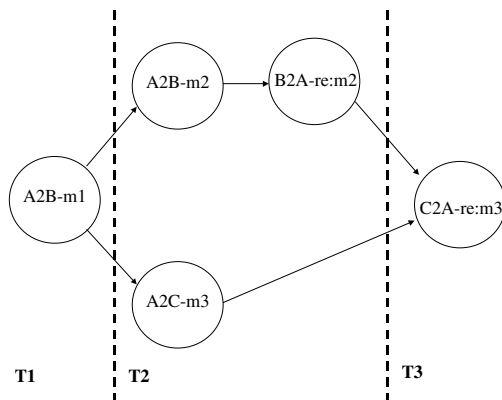


Figure 1: Interaction process at time T3

At a given time t , the model must provide a precise description of the history of the transaction up until t . The model can also provide indications of possible developments of the transaction. Considering the example in (Figure 1), at the end of time T2 the model must provide a representation of the first four interactions. An interaction node X2Y-m indicates that X has sent the message m to Y. The model can indicate that at a future time C might send a reply to A for m3, possibly indicating a degree of confidence. History is essential. Prediction is desirable.

The information model we propose revolves around three capabilities: aggregation, structuring, and prediction. The techniques adopted for the enforcement of the model exploit the peculiarities of B2B interaction. The structure of a typical business message is presented in (Figure 2).

3. Aggregation of the data

Given the overall flow of business messages, the first level of aggregation is based on the set of business partners involved. One-to-one is the most common form of business interaction, but multiparty is also quite important. Information on the business partners involved in a message can be extracted from the headers of a message. The second level of aggregation is based on the business transaction to which the message is related. Different parts of a company can use different processes and systems. However, the concept of transaction as unifying element for business activities is well established at business level. Some form of identification for business transactions is always present in business communication. In order to route efficiently the message to the appropriate business entity, information are usually present also on the type of both the message and the transaction. For example, the message confirming the payment of an invoice contains indications on order number, invoice

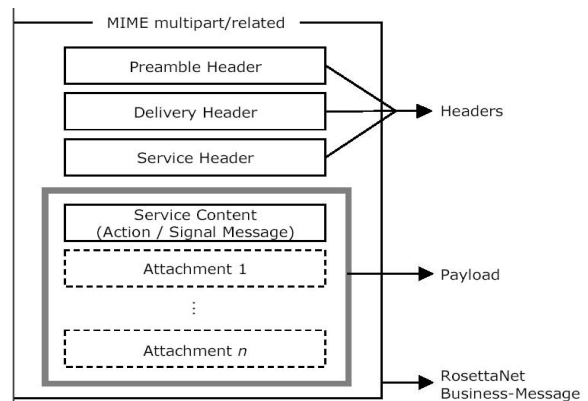


Figure 2: RosettaNet business message

number, and the indication that the message relates to the financial aspect of a transaction.

From a technical perspective, the data needed for the level of aggregation we require is accessible from the headers and the content descriptor of messages. Such data are exposed in a structured format, and can be accessed quite efficiently. Other parts of the message can be subject to some form of application-level encryption. Headers and service description are usually sent in clear, or subject only to transport-level encryption (e.g. SSL).

4. Process-oriented structuring of data

As the messages pertaining to a specific transaction accumulate, the information model requires a specific structure to be imposed on the data. Causal and temporal dependencies are the key elements of the structure required. The goal is to reconstruct on-the-fly an a-posteriori description of the interaction process between the parties. New messages monotonically extend the process structure.

Temporal dependencies have been modelled using the basic concept of time intervals (Figure 1). Information on the time at which a message is sent or received can be derived from transport metadata. All the messages observed in a given time interval are considered part of parallel threads, unless causal dependencies are identified. Messages observed in distinct time intervals are assumed sequential. The sequential relation is transitive, and it can be represented in an intentional form. All the parties require a consistent view of the observation time for a message, hence standard solutions are put in place to avoid ambiguity on border cases.

The management of causal dependences is based on the message type and on specific knowledge on message flows. A knowledge base contains the general rules describing the dependencies at type level. In addition, instance-specific information is used. For example, a

general rule can state that a payment message is followed by an acknowledgment of a given type. Still, there can be multiple payments within the scope of a specific transaction. Type-specific and instance-specific rules are used at the same time. From a technical perspective, the level of information exposed by the message structure allows a rich set of rules to be inserted in the knowledge base. In practice, message flows reflect the structural and efficiency-driven simplicity of business processes. For example, RosettaNet PIPs [4] provide correlation information for atomic interaction in cooperative processes.

5. Flow prediction

Predictions on the future directions of a message flow provide useful information for management and optimisations. The type of information we use derives in part from the same knowledge base used for causal correlation and in part from statistical information. Both types of information are represented in the form of process patterns. From an operational perspective, we apply pattern-matching techniques to process branches. The order of application goes from the most recent leaves towards the root. Matches from the static knowledge base have precedence over statistical information.

The objective of the prediction technique is to indicate likely developments for an interaction process. Multiple possibilities can be explored at the same time. Probabilistic concepts are embedded into the process model, and statistical information is accumulated in the knowledge base. The information on the actual development observed for the process contributes to the continuous refinement of the information the knowledge base used for the predictions.

6. Summary

E-business processes play for the supply chain the role that internal processes play for business operations. Both from a business and from a technical perspective, e-business processes have a number of commonalities with internal processes. Still, the peculiarities of e-business processes require a new approach in terms of modelling techniques as well as a management infrastructure. The intrinsic capability to adapt and evolve represents a major requirement for e-business processes.

In this paper, we propose an approach to the modelling and management of e-business processes based on on-the-fly inference of procedural interaction logic. The techniques proposed rely on the direct observation of the flow of business messages generated by a company. Explicit upfront knowledge is combined with implicit knowledge emerging from the message flow. RosettaNet is used as reference standard for business messages.

References

- [1] AA.VV. *"The third imperative of e-business. The role of collaboration networks in e-business"* CAP Ventures, 2000.
- [2] Holligsworth D. *"The workflow reference model"* Workflow Management Coalition (WfMC), TC00-1003, 1994.
- [3] Piccinelli G. and Stammers E. *"From e-processes to e-networks: an e-service-oriented approach"* Proc. Workshop on Object-Oriented Web Services at OOPSLA, Tampa, Florida, USA., 2001.
- [4] RosettaNet, Documentation for the Standard, <http://www.RosettaNet.org>
- [5] Seaborne A., Stammers E., Casati F., Piccinelli G., and Shan M. "A framework for business composition" In Proc. *W3C Workshop on Web Services*, San Jose, CA, USA, 2001.