

CRM System Using PA-AKD Approach of D³M

Ajay kumar^{#1} A.Tejaswi^{\$2} Lakshmi Prasad Koyi^{&3} G.Narasinga Rao^{\$4}
S.Srinivasa Rao Illapu^{*5}

^{#1} Asst.Prof, Dept of CSE/IT, Dronacharya College of Engg, Greater Noida.
U.P, ajaykeshri@gmail.com.

^{\$2} Dept. of Information Technology, GITAM University, Visakhapatnam,
tejaswiorama@gmail.com.

^{&3} Asst. Prof, Dept of CSE, CMR Institute Of Technology, Hyderabad.
prasad.koyi@gmail.com.

^{\$4} Asst. Prof, Dept of CSE, GMRIT, Rajam, narasingarao.g@gmrit.org.

^{*5} Asst. Prof, Chaitanya Engineering College, Visakhapatnam,
srinivas_sankar2001@yahoo.co.in.

ABSTRACT

In today's business world, Customer Relationship Management has become conventional. In spite of the unpredictable economy, CRM is being forced into corporate budgets and is believed to be the leading and initiative factor by many companies. The custom of CRM system has become mandatory. An adequate CRM system helps in acquiring new customers and increasing the value of existing customers. Data warehouse acts as the data resource for D³M on CRM. This paper introduces the CRM based on PA-AKD and managing relationships with customers which has become the critical competition between organizations. Then the structure of CRM is described with the help of D³M, discusses the procedure to Data warehouse and the significance of D³M applied to CRM. At the end, the promoting effects of D³M to individual service are presented.

KEYWORDS

CRM, Domain Driven Data Mining (D³M), Data warehouse, AKD, E-commerce.

1. INTRODUCTION

Electronic commerce (or E-commerce) encompasses all business conducted by means of computer networks. Advances in telecommunications and computer technologies in recent years have made computer networks an integral part of the economic infrastructure. More and more enterprises are facilitating transactions over web. So enterprises recognize the importance of improving their ability to manage customer effectively. E-commerce provides multiple benefits to the consumers in form of availability of goods at lower cost, wider choice and saves time. Customer has become the decisive main resource [4]. Therefore, Customer Relationship Management (CRM) is well considered as one effective way to keep the competitive advantage and the level of CRM is the key for winning competition.

CRM is a set of strategies, processes, metrics, organizational culture and technology solutions that enhance an organization's ability to see the differences in its customers' and prospects' behavior and needs, track new opportunities to better serve their customers and act, instantly and profitably, on those differences and opportunities. The aim of CRM is optimize the use of technology and human resources for the business to gain insight into the behavior of customer. Recently CRM has taken a center stage in the business world with businesses concentrating on saving money and increasing profits by redefining internal processes and procedures. It costs a company dramatically less to retain and grow an existing client, than it does to court new ones. It is said that "It is seven times more expensive to acquire a new customer than to keep an existing one". It utilizes the modern communication technology, with Business Process Reengineering (BPR) [7], increases enterprise's ability of retaining the former customers and finding the new customers, and maximizes the benefit from customers. CRM includes total

process which is composed of judging, selecting, expanding, attracting and retaining customers. With the development of the market economy, more and more companies are conscious that it is not enough to only depend on high quality products and sales network to get customers. They have to be concerned about the requirements of customers, increase the response speed and provide the "one to one" service for customers. Then they can acquire the trust of customers and prevail in market competition [3].

CRM systems enable enterprises become active from passive in services. CRM systems on net can help enterprises to store, organize and analyze the essential information of customers, to customize the service item data and correlated history data. According to the information of customers, enterprises can provide with overall, timely and personalized service for customers [5]. At the same time, enterprises can find out the new requirements of customers and create new business increases. But when designing and developing CRM systems of the Browser/Server architecture, there are the following problems in acquiring information efficiently. Because the running platform of CRM systems is based on Internet and the information is no longer localized only on LAN, so the acquiring information problem must be solved effectively. With the development of advanced computer technology, D³M becomes a method or tool that can aid companies to become more customer-oriented. Moreover, the customer life cycle of CRM consists of three phase's customer acquisition, customer relationship enhancements and customer retention. The customer lifecycle provides a good framework for applying D³M to CRM. On the "input" side of data mining, the customer lifecycle tells what information is available. On the "output" side, the customer lifecycle tells what is likely to be interesting.

Most data mining algorithms and tools stop at the mining and delivery of patterns satisfying expected technical interestingness. There are often many patterns mined but business people either are not interested in them or do not know what follow-up actions to take to support their business decisions. This issue has seriously affected the widespread employment of advanced data mining techniques in greatly promoting enterprise operational quality and productivity. Domain Driven Data Mining (D³M for short) targets the development of next generation data mining methodologies, frameworks, algorithms and decision support, which aim to promote the paradigm shift from data-centered hidden pattern mining to domain-driven actionable knowledge discovery(AKD) As a result of the D³M research and development AKD system can deliver business-friendly and decision-making rules. D³M can bring about the effective and practical development of many challenging data mining applications in every area [1].

2. ROLE OF AKD IN D³M

In the real world, data mining is a problem-solving process (R) from business problems Ψ (with problem status τ) to problem-solving solutions Φ

$$R : \Psi(\tau_1) \rightarrow \Phi(\tau_2)$$

Gradually, data miners realize that the actionability of a discovered pattern must be assessed by and satisfies domain user needs. To achieve business expectations, business interestingness measures to what degree a pattern is of interest to a business person from social, economic, personal and psychoanalytic factors. Recently business objective interest is recognized by some researchers, say profit mining and domain-driven data mining, involving business interests. Moreover, business subjective interest also plays essential roles in assessing business interests. This leads to a comprehensive cognition of actionability. There are two sets of interest measures needed to be calculated when a pattern is extracted. For instance, we say a mined association trading rule is (technically) interesting because it satisfies requests on support and confidence.

In the real-world mining, business interests may differ or conflict technical significance. Clearly, actionable knowledge mining targets patterns confirming the relationship that the pattern satisfies both business expectations as well as technical significance. However, it is a kind of artwork to tune thresholds and balance significance and difference between technical and business interestingness [1].

There are two steps in technical interest evolution. The original focus basically was on technical objective interest, which aims to capture the complexities of pattern structure and statistical significance. Technical subjective measures, which also recognize to what extent a pattern is of interest to a particular user.

AKD (Actionable Knowledge Discovery) is an iterative optimization process toward the actionable pattern set P, considering surrounding business environment and problem states. From the modeling perspective, an AKD-based problem-solving process is a state transformation from the source data DB($\Psi \rightarrow DB$) to the resulting pattern set P($\Phi \rightarrow P$). Let DB be a database related to business problems Ψ , $X=\{x_1, x_2, \dots, x_L\}$ be the set of items in DB, where $x_l (l=1, \dots, L)$ be an itemset, and the number of attributes in DB be S. Suppose $E = \{e_1, e_2, \dots, e_K\}$ denotes the environment set, where e_k represents a particular environment setting for AKD. Further, let $M = \{m_1, m_2, \dots, m_N\}$ be the data mining method set, where $m_n (n = 1, \dots, N)$ is a method. For method m_n , suppose it is identified pattern set $P^{mn} = \{P_1^{mn}, P_2^{mn}, \dots, P_U^{mn}\}$ includes all patterns discovered in DB, where $P_u^{mn} (u=1, \dots, U)$ is a pattern discovered by m_n [1].

$$\begin{array}{c}
 \text{AKD} \xrightarrow{e, \tau, m \text{CM}} \text{DB} \xrightarrow{e, \tau, m_n} P^{mn} \\
 \longrightarrow \text{O}_{pCP}^{e, \tau, m \text{CM}} \text{Int}(P) \\
 \longrightarrow \tilde{P}
 \end{array}$$

AKD is critical in promoting the productivity of data mining and knowledge discovery for smart business operations and decision-making rules. With regard to AKD approach, the existing work mainly focuses on developing post-analysis techniques to filter/prune rules, reduce redundancy and summarize learned rules. Real-world data mining is a complex problem-solving system. From the view of systems and micro economy, the endogenous character of AKD determines that it is an optimization problem with certain objectives under a particular environment [1].

3. ARCHITECTURE OF CRM BASED ON PA-AKD FRAMEWORK

PA-AKD is a two-step pattern extraction and refinement exercise. First, generally interesting patterns (which we call ‘general patterns’) are mined from data sets in terms of technical interestingness ($t_o()$, $t_s()$) are used. Further, the mined general patterns are pruned, distilled and summarized into operable business rules (embedding actions) (which we call ‘deliverables’) in terms of domain specific business interestingness ($b_o()$, $b_s()$) and involving domain (Ω_d) and meta (Ω_m) knowledge. PA-AKD is a two-step optimization problem that can be expressed as follows [2]

$$\text{PA-AKD: } \text{DB} \xrightarrow{e, t_i(), m_1} P \xrightarrow{e, b_i(), m_2, \Omega_d, \Omega_m} P, R$$

In Post Analysis, a recent highlight is to extract actions from learned rules. A typical effort on learning action rules is to split attributes into ‘hard/soft’ or ‘stable/ flexible’ to extract actions that may improve the loyalty or profitability of customers. The existing post-analysis and post-mining focuses on association rules or their combination with some specific methods. This limits the actionability of learned actions and the generalization of proposed approaches for AKD. For a pattern p, Int(p) can be further measured in terms of technical interestingness ($t_i(p)$) and business interestingness ($b_i(p)$).

$$Int(p) = I(t_i(p), b_i(p))$$

The interestingness system, which combines technical interestingness ($t_i()$) with business expectations ($b_i()$) into a Post Analysis AKD interestingness system ($i()$). Domain knowledge (Ω_d) and Environment (e) must be considered in the data mining process. Finally the outputs are P and R. Correspondingly, the actionability of a pattern p is measured by $act(p)$ [2]:

$$\begin{aligned} act(p) &= O_{p \in P}(Int(p)) \\ &\rightarrow O(\alpha t_0(p)) + O(\beta t_s(p)) + \\ &\quad O(\gamma b_0(p)) + O(\delta b_s(p)) \\ &\rightarrow t_0^{act} + t_s^{act} + b_0^{act} + b_s^{act} \\ &\rightarrow t_i^{act} + b_i^{act} \end{aligned}$$

$O(.)$ is the optimization function to extract function to extract those $p \in P$, where $Int(\tilde{p})$ can beat a given benchmark. Whereas t_0^{act} , t_s^{act} , b_0^{act} and b_s^{act} measure the respective actionable performance in terms of each interestingness element. Due to the inconsistency often existing in different aspects, we often find identified patterns only fitting in one of the following sub-sets:

$$\begin{aligned} Int(p) \rightarrow &\{ \{ t_i^{act}, b_i^{act} \}, \{ \neg t_i^{act}, b_i^{act} \}, \\ &\{ t_i^{act}, \neg b_i^{act} \}, \{ \neg t_i^{act}, \neg b_i^{act} \} \} \end{aligned}$$

where ' \neg ' indicates satisfactory [1, 2].

The architecture of CRM system based on PA-AKD framework in data warehouse as shown in Figure1.

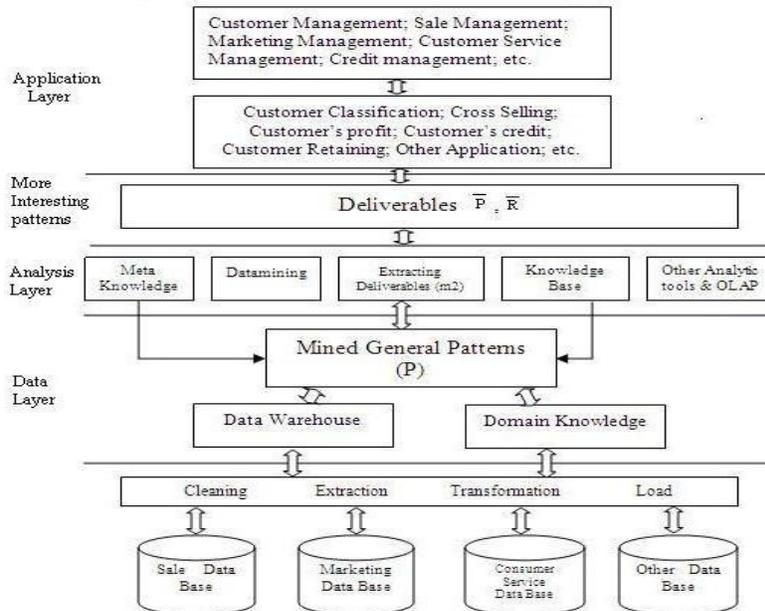


Figure 1 The architecture of CRM

The data has been collected from various databases of consumer sales, services and marketing etc. These various databases are processed through cleaning, extraction, transformation and loading techniques of Knowledge Discovery in Databases (KDD). The metamorphosis data are accumulated into the Data warehouse and Domain knowledge. The mined general patterns of business intelligence $b_i()$ and technical significance $t_i()$ are extracted from both the analysis and data layer. Deliverables P and R are analyzed and extracted from the mined general patterns.

These deliverables are used for the customer services and management in the application layer. The deliverables of PA-AKD can be efficiently used to CRM system.

The key point in this framework is to utilize both domain/meta knowledge and business interestingness in post-processing the learned patterns. In the real world, this framework can be further instantiated into varied mutations [9], [10], [11]. In fact, many existing methods, such as pruning redundant patterns, summarizing and aggregating patterns to reduce the quantity of patterns, and constructing actions on top of learned patterns, can be further enhanced by expanding the PA-AKD framework and introducing business interestingness and domain/meta knowledge into the AKD process. [12]

4. DESIGN AND ACTUALIZATION OF THE DATA WAREHOUSE

Setting up data warehouse system is a complex process, the key of which is how to transform data from application oriented structure to analysis oriented data structure, and can combine closely with the actual requirement to meet the analysis of decision-making. In this system, although the data amount needing deal with is relatively great, but the type of data is monotonous, data structure is relatively simple. So we choose Microsoft SQL Server 2000 as developing tool of data warehouse, and Microsoft SQL Server Analysis Services as online analyze process and data mining tool, use English Query as SQL syntax transform and English Query tool [3].

4.1 Data Modeling

CRM systems usually work with large sets of data and require a short response time and if you consider using analytical tools like OLAP against virtual data warehouses then you have to build your system through SQL queries and retrieve data directly from OLTP systems. In this case, the large amount of data in ERP systems may lead to an increase of responding time for BIS. That's why you should consider applying optimization techniques in order to improve the BI system's performance. Metadata of data warehouse could be created from value of gathering data calculating data in severally time-sequence. After the collection of the customer's information, we will convert each parameter into the time-sequence parameter. Dimension of data warehouse are cube which is made up of time and values of deflection. Then, the data warehouse can be constructed according to our need [3].

4.2 Data Extraction

Extraction is the operation of extracting data from source systems for further use in a data warehouse environment. After the extraction, this data can be transformed and loaded into the data warehouse. Designing and creating the extraction process is often one of the most time-consuming tasks and, indeed, in the entire data warehousing process. The data has to be extracted normally not only once, but several times in a periodic manner to supply all changed data to the warehouse and keep it up-to-date. Moreover, the CRM system typically cannot be modified, nor can its performance or availability be adjusted, to accommodate the needs of the data warehouse extraction process [3].

4.3 D³M and Knowledge Discovering

D³M aims to construct next-generation methodologies, techniques and tools for a possible paradigm shift from data centered hidden pattern mining to domain-driven actionable knowledge delivery. According to the architecture designed, using Microsoft SQL Server Analysis Services as tools obtains the fuzzy rule of control rules by carrying on the pretreatment, converting, and mining, explaining to the data and new rule can be picked up continuously, and the system rule database can be revised online. In this way, the content in the

rule database of system will be changed constantly as time goes on, make the system more meet demand of actual production process [3].

5. ADOPTION OF D³M TO CRM

According to an enterprise view this CRM helps an enterprise to enable its marketing departments to identify and target their best customers, manage marketing campaigns with clear goals and objectives, and generate quality leads for the sales team. Allowing the formation of individualized relationships with customers, with the aim of improving customer satisfaction and maximizing profits; identifying the most profitable customers and providing them the highest level of service. Providing employees with the information and processes necessary to know their customers, understand their needs, and effectively build relationships between the company, its customer base, and distribution partners. CRM is being thrust into corporate budgets. It has gone from being an important edge in the business world to a necessary tool for survival. CRM in its broadest sense simply means managing all customer interactions. In practice, this requires using information about enterprises, customers and prospects to more effectively interact with their customers in all stages of relationship with them. We refer to these stages as the customer life cycle [3]: acquiring customer, increasing the value of the customer and retaining good customers. Domain driven Data mining can improve efficiency in each of these stages through CRM system.

5.1 Attracting New Customers

The major indexes of business development include the ability of acquiring new customers. The first step in CRM is to identify prospects and convert them to customers. And this can be achieved by activating the potential demand of customers that know little about your products, and make them produce behavior of buying to become enterprises customers. CRM provides employees with the information and processes necessary to know their customers, understand their needs, and effectively build relationships between the company, its customer base, and distribution partners. It also allows the formation of individualized relationships with customers, with the aim of improving customer satisfaction and maximizing profits; identifying the most profitable customers and providing them the highest level of service. The first step in CRM is to mine the customer's as many feature description as possible [3]. On one hand, in the process of enterprise's informationization, marketing facing the situation has changed from a traditional mass of readers and direct group to passed network marketing and E-mailing etc. On the other hand, enterprises and trade companies collect the consumer behavior information of a large number of customers through various kinds of ways and classify the customer according to the set indexes. Then looks forward for the masses of potential customer's selling point and follow relevant business activities thus carry on marketing pointedly, in order to activate potential consumption, make the potential customer change into the real customer. Contact management module and direct marketing module of CRM allow companies to effectively promote and market their products and services to prospects. Those modules help speed up the acquiring processes and reduce the cost of acquiring new customers.

5.2 Quantifying the Value of Qualitized Existing Customers

CRM helps companies better understand existing customers' needs and behaviors and enhance the profitability from existing customers by cross selling. They can customize their products and services to individual customer's needs and preferences. The relation between modern enterprise and customer is often changed. In order to make this kind of customer's relation become more and more perfect, cross-selling need to be carried to the existing customer. The customer gets more easy service of meeting his demand and benefits from it, enterprises make a profit because of growth of the sales amount. Through cross-selling, promote the existing customer's profit ability [3].

The D³M can realize cross-selling to excavate effectively marketing, offer new products and service to the existing customer. It supports customer's profitability to analyze and predict customer's profitability change. To the existing value customer, enterprise can adopt individualized marketing and service tactics, increase customer's loyalty to the old brand, make them become new supporters and followers of brand at the same time to strive to make. The domain-driven data mining is excavated to find about these existing customers easily by the enterprises. A large amount of data can strengthen the accuracy that the data excavate to support [3].

5.3 Retaining Good Customers

Customer retention is critical to the overall profitability of an organization. A customer you spend hundreds of dollars and months to acquire may leave you in seconds as a result of poor customer services. The cost that enterprises obtain the new customer is rising constantly. A lot of companies agree that the cost of getting one new customer is 6-8 times of cost to retain an existing customer [3]. We can divide a large number of customers into 3 kinds according to attribute with data classification in data mining. The first class is a valueless customer. The second class is a steady valuable customer. The third class is unstable to look for favorable price and fine valuable customer of service constantly. Customer relation management thinks that the third class customer is most important to be retained, and it will reduce enterprise's operation cost to do it in this way.

6. D³M ACCESSIBILITY TO INDIVIDUAL SERVICE

D³M for service in individuation is a kind of service modes that supplies different services depending on different customers. It is the best choice when compared to traditional service modes. Individual service on CRM is to supply different services depending on different customers and new application and development of service in individuation [6]. With development of E-commerce, customers need more time and energy to find something interesting and the great quantity of information on internet rather than they do in traditional commerce. In this case, customers need service modes that automatically organize and adjust information according to their demands.

D³M that is the most important step to find knowledge in E-commerce consists of two parts: one is to mine information which customers visit web sites to find behavior and mode which customers browse web sites so that customers' interests can be found; the other is to mine data which customers exchange to find association relation between goods so that customer' latent needs can be found. Some methods of D³M can promote individual service even customer relationship management. Many research issues of D³M for discovering knowledge in E-commerce that effectively synthesize the above ubiquitous intelligence in AKD-based problem-solving systems are listed below [1, 8].

6.1 Data Intelligence

It includes mining in-depth data patterns, and mining structured knowledge in unstructured data. It mines main characteristics from a special group with some same attributes. These characteristics can be used to produce new data items about customers added to database [1, 8].

6.2 Domain Intelligence

It assists in understanding and problem-solving of the problem. Domain intelligence consists of qualitative and quantitative intelligence. Both types of intelligence are instantiated in terms of aspects such as domain knowledge, background information and organizational factors [1, 8].

6.3 Network Intelligence

It comprises information retrieval, text mining, and web mining. By the method, customer characteristics visiting a commerce web site can be found according to statistic information and visiting mode. So latent customers can be found and effective commerce measures can be taken [1, 8].

6.4 Human Intelligence

This refers to (1) explicit or direct involvement of humans such as empirical knowledge (2) implicit or indirect involvement of human intelligence such as imaginary thinking, emotional intelligence. We can get a similar customer group according to information in servers. That is to say, it can produce items into a same set explicitly. By the method, marketing strategies can be improved in E-commerce. Automatically send sales mails to special customers implicitly, and when any customer from different groups visit a web site, the web can change its pages for him or her, for example. By the measures, demands of customers can be met in some way to reach marketing targets [1, 8].

6.5 Social Intelligence

It includes collective intelligence, social network analysis, and social cognition interaction. It not only gives a theory frame, but also helps to manage goods and improve CRM services in E-commerce [1, 8].

6.6 Intelligence Metasynthesis

The above ubiquitous intelligence has to be combined for the problem solving. The methodology for combining such intelligence is called metasynthesis, which provides a human-centered and human machine cooperated problem solving process by involving, synthesizing and using ubiquitous intelligence surrounding AKD as need for problem solving. By the technology, on the one hand, association between pages files visited by customers in a session can be found by mining logs in a web site. On the other hand, goods can be also found by mining trade transaction databases. These will be helpful to any E-commerce web site to organize page structures and make strategies of marketing [1, 8].

7. CONCLUSION

CRM, a strategy that influence very advanced technologies in such a way that one can give a good competition in business world of 21st century. If the organizations are able to use the information about enterprise customers and meet their needs then those organizations can come to a successful position. The means of using CRM is to learn more about customer's needs and behaviors in-order to maintain and develop stronger relationships with them, CRM by using D³M is the essential guide.

Infact, the D³M is closely linked to the large potentiality of e-commerce with future of CRM. Customer Relationships can be improved by technical interestingness and business expectation patterns. PA-AKD system is used to extract interesting patterns from the database of the customer which results in more extensive application prospect and market of CRM. The value must make the enterprise a different level competer and play a vital role in core competition which strengthens the enterprise.

REFERENCES

- [1] Longbing Cao, (2008) "Domain Driven Data Mining (*D3M*)", IEEE International Conference on Data Mining Workshops.

- [2] Longbing Cao, Yanchang Zhao, Huaifeng Zhang, Dan Luo, Chengqi Zhang, E.K. Park, (2008 Jan) "Flexible Frameworks for Actionable Knowledge Discovery", *Journal of Latex class files*, vol. 1, no. 1.
- [3] Dezhen Feng, Zaimei Zhang, Fang Zhou, Jianheng Ji, (2008) "Application Study of Data Mining on Customer Relationship Management in E-commerce".
- [4] B.Zhang, K. Li. , (2004) "Customer Value Evaluation and Management Model Based on Cost.", *Finance Communication*, (09):7-12.
- [5] R.Han, Z. Tian. (2003) Evaluation Method of Customer Value, *value Engineering*, 36(06):15-16.
- [6] C.Zeng, X. Xing., (2002) "A Summarization of Service in Individuation", *Journal of Software*,20(10):63-67.
- [7] C.Zhang (2002) " Analysis on Customer Value Management, *Science and Technology*", 84(02):4-15.
- [8] Longbing CAO (2009), Domain Driven Data Mining, PAKDD2009 Tutorial, Bangkok, Thailand, lbcao@it.uts.edu.au, University of Technology Sydney, Australia.
- [9]. B. Liu, W. Hsu, and Y. Ma, (1999) "Pruning and Summarizing the Discovered Associations", in *SIGKDD*.
- [10]. B. Liu and W. Hsu, (1996) "Post-Analysis of Learned Rules," in *AAAI/IAAI*, 1996.
- [11]. Q. Yang, J. Yin, C. Ling, and R. Pan, (2007) "Extracting Actionable Knowledge from Decision Trees," in *IEEE Transactions on Knowledge and Data Engineering*, 19(1):43-56.
- [12]. L. Cao, "Developing Actionable Trading Strategies,(2008) " *Knowledge Processing and Decision Making in Agent-Based Systems*, N. Nguyen and L. Jain, eds. Springer.