A Scope of Chance Discovery

Yukio Ohsawa

Graduate School of Business Sciences, The University of Tsukuba, Japan 112-0012

Abstract

What we call *chance discovery* is the discovery *of* a chance, not *by* chance. Here, a ``chance'' is defined as a piece of information, about an event or a situation which gives significant impact on the decision-making of human, agents, and robots. This follows the second meaning of ``chance''- *a suitable time or occasion to do something* - in the Oxford Advanced Learner's Dictionary, considering that the situation or event occurred at a certain time is more significant than the time itself, and that a decision to do something precedes the action of doing it. We defined chance discovery as the awareness of a *chance*, defined above, and the explanation of its significance. This triggered researchers in various domains to challenge exploring methods for chance discovery even if the chance is rare and its significance has been unnoticed. They are going much beyond previous data mining and statistics by introducing various ideas.

Which do you want, a rule or an opportunity?

Do you like a rule

"if it is Sunday morning, one must go out for jogging"

? When I asked this question to colleagues, some disliked jogging because of its heavy feeling. Some rather preferred a rest on holidays, and others wanted to decide what to do flexibly reacting to the dynamic environment e.g., the weather. They look out of the window to see the weather condition, or watch weather forecast programs on TV.

In the recent business literatures, the word "knowledge" has been increasing its frequency of appearance. For example, "knowledge management" and "knowledge discovery" are the most popular terms bundled with keywords on information technology. However, the definition of "knowledge" has been hardly consistent among researchers. In the literatures of management, knowledge has been distinguished into implicit (not symbolized, although existing in the mind of human) knowledge and explicit (symbolized) knowledge. In the case above of jogging on Sunday, what kind of knowledge should be activated for deciding to stay home, when we look outdoors? A person who decided may answer:

"I used the explicit knowledge that one should stay home when it is raining." Is this a valid knowledge? No, she could have used other knowledge, e.g.,

"I go out even if raining, as far as it is not thundering."

Each piece of knowledge used by human in real situations is a part of complex knowledge with a large number of conditions (i.e., the terms in the "IF …" part), if the knowledge can be symbolized as an IF-THEN rule. Rather than believing in her comment, it would be more realistic to say she had experiences of weather-based actions, and decides her own action based on these experiences and the on-site information of the current weather. In other words, she likes to act reactively with changes in the environment than to be bound by a rule much simpler than the dynamics of the real world even if it has one thousand conditions. In this article, let us assume situations where human seek a chance, i.e., information about a new event or a situation significant for making decisions in face of the dynamic environment..

Is Chance Discovery itself an Event or a Process?

The word "chance," means an event or a situation significant for the decision of a human, an agent, or a robot. We need at least 10 minutes to evaluate a pattern "event B occurs within 10 minutes after event A" whereas each event A and B may occur at one moment. What about the discovery of a chance? Is it an event to start and end within a moment?

Here let us try to answer a question, "Can a chance be discovered in a moment, if the chance occurs at a moment?" "Yes," one may answer, "when I hear a new message, I can immediately distinguish a good news from a bad news. Just think of weather news – you do not take long before feeling happy after you hear it will be fine all the day…"

We have a reason to reject his answer. Why does he feel happy at the moment he catches the news? He feels so because he has been desiring a fine day, because of several possible reasons, e.g., he was planning to go out or he likes the blue color of sky by nature. We can hardly show an example where one can feel an opportunistic event occurred, without one's preparatory state of mind. In other words, *chances come to a prepared mind.*

All in all, the discovery of a chance by human is realized as a process including his/her *involvement* in a dynamic environment. Even in such a situation, human must continue to make certain decisions (sometimes s/he decides to do nothing) and keep feeling he needs essential pieces of information on which to making decisions. For example, one will strengthen one's attention to weather news during a season while the weather keeps changing. That is, one's *concern* with chances increases through this involvement in the dynamics of the surroundings. In this phase, one comes to be ready to catch the information about a new occurrence of event as his or her chance. Being *aware* of the (occurrence of) chance, i.e., paying attention to the chance strongly enough, one can begin to understand its meaning into details. Finally, one can decide to take advantage of the chance for making a decision. Thus, chance discovery is likely to be a process rather than a momentary event.

Many problems still remains. For example, does the awareness of a chance always precede understanding it? We cannot say "yes" simply, because we usually cannot pay attention to a new event feeling its significance, if we do not have the slightest understanding about the event. Chance discovery is a complex process in which the occurrence, the awareness, and the understanding of chances interleave each other. This section is dedicated to illustrate the process of chance discovery, where human and computer collaborate for mining chances from dynamic environment.

From Knowledge to Chances

Complex rules have been learned for predicting a predefined rare event, such as a break of signal transmission at a certain line, by the extension of genetic algorithm, cost-sensitive learning, etc (Weiss and Hirsh, 1998). These work for automated prediction where the candidates of prediction target, i.e., significant rare events, are not *selected* on-line by the prediction system but are *predefined* off-line. That is, the obtained rules are optimized according to a certain predefined object function corresponding to the likelihood of the occurrence of the target event under the observed conditions. All in all, by such existing data analysis/mining methods we can not explain the significance of (1) unnoticed features not described in the given data, or of (2) rare events (values of existing features) people seldom count as worth while predicting.

In extracting and explaining significant financial anomalies and their causes, new products and their latent popularity, and the signs of fatal hazards and their meaning to human life, we always suffer from both difficulties of (1) and (2) above. For example, in marketing, the emotion of people may form values (the strength of emotion) of features (anger, desire, etc.), and their desire for neat clothes may form a new market for cloth companies.

Keys to Open Ways to Chance Discovery

For evaluating the significance of a new factor (a feature or its value corresponding to (1) and (2) above) unfamiliar to usual people, it becomes necessary to shift one's mental context from usual situation to be concerned with unusual situations where the factor becomes influential. Otherwise human cannot smoothly accept a new situation, because of the disturbance by cognitive dissonance (Festinger, 1957) due to the unfamiliarity with the situation. On the notion of "context shifting," the process of chance perception has been modeled (Nara and Ohsawa, 2000). It is essential to shift the mental context of most people who are not concerned with hidden factors potentially causing new events, to accepting the new situation and to reasoning whether the event is desired or scared of - as a chance.

For the early stage of context-shifting, the essential information is *peripheral* (Petty and Caccioppo, 1981) as in Fig.2, i.e. information not directly explaining the significance of the chance but attracting people to chances. For example, a persuasive authority (e.g., a renowned professor of seismology) showing a visualized simulation where a certain event becomes a meaningful chance (a small quake preceding a big earthquake, displayed on TV) works as a peripheral clue. If the subject accepts the new context, *central information* directly explaining the significance of the chance becomes acceptable and pushes her forward to a real *action*. This central information is sometimes acquired from recent data of event occurrences. Possibly, a familiar part of the data may attract user's attention to the peripheral environment of the chance, as put "peripheral clue 2" in Fig.2. During real actions, she can evaluate the current chance and triggers herself to desire the next chance if necessary. It is noteworthy, however, that spiral models of chance perception (Shibata and Hori, 2001; Ohsawa, 2002) commonly point out that an action may concentrate the attention of the subject too deeply into the problem in face, to become aware of other new chances.

The process in Fig.2 distinguishes chance discovery from knowledge discovery from data (KDD) (Fayyad et al, 1996) in Fig.1 in essential aspects. The ``interpretation'' in Fig.2 is separated from "evaluation" reflecting the discussion above. The spiral model of human perception of chances is integrated with the KDD process and human is positioned as a stronger engine than machine to select, interpret and evaluate the value of a chance in Fig.2. This evolution of the process model directs the studies on chance discovery to answering the question "How can we human discover chances?" not leaving chance discovery only to the power of computer.

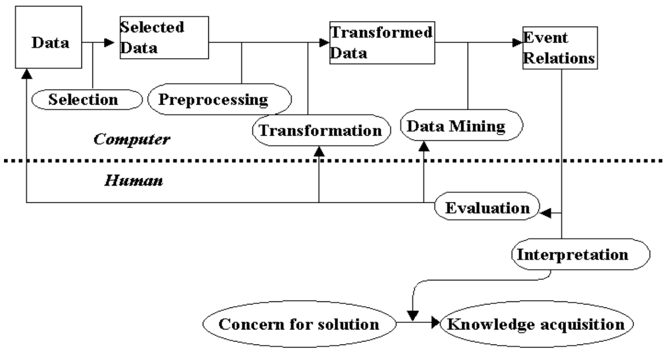

Fig. 1 Fayyad's model of knowledge discovery process.

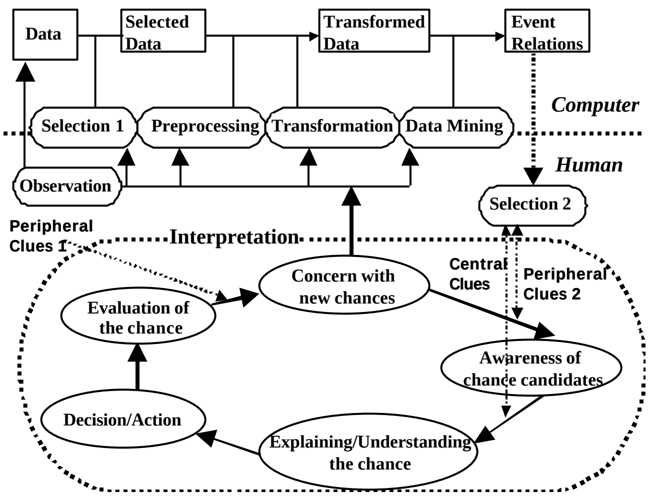

Fig. 2 The process of chance discovery

A chance is here evaluated on the value of decisions or actions, interacting with external environment, based on the chance. In this interaction, new essential features are detected and reflected onto the ``observation'' of the value of those features.

The contribution of human to the ``selection'' step, where rare events should be picked if they may really be chances, is entrusted to the interpretation cycles where the interpretations of significant chances are recognized as ``central clues'' (``Selection 2''). The "selection" as was in the KDD process i.e., ``Selection 1'' in chance discovery plays a rather weaker role in chance discovery, discarding only obvious noises (e.g., ``is,'' ``are,'' etc. in the case of text data). In the interpretation cycles, as in the cycles of SECI (Nonaka, 1995) going through implicit and explicit communications among a group of co-working people, people may reach an explicit understanding of chances by selecting and diffusing the most useful interpretations of the chance among opinions from a few participants.

Real data can be an efficient external information source to guide human to the explicit awareness and explanation of a chance, if shown in a suitable manner. For example, by a data mining method showing causalities including a rare event, a human in a ``concerned'' state is likely to understand the central explanation about the significance of rare events Hereby, we noticed three keys to chance discovery:

Key 1: Communications about the central information (i.e., the interpretations) of the significance of an event.

Key 2: Peripheral information (e.g., high-impact scenes visualizing the chances, words from authorized people, etc) for enhancing user's concern with (or imagination of) the context where, i.e. the reason why, a chance is desired.

Key 3: New methods of modeling and mining chances from data, for obtaining and showing the causalities involving rare events.

Let us go into basic theories and approaches to Key 1 through 3, in the following.

Communications, Key 1 to Chance Discovery

People share some common senses about daily affairs, but can have different knowledge about new situations. External information sources such as books and Web pages may supply new knowledge, but human is rather accustomed to consult suitable people for advices on what to do in confronted situations. Especially, when the situation is new and uncertain, human-human communication becomes essential for noticing, diffusing and socially establishing useful knowledge (Allport and Postman, 1947).

Let us look briefly at the decision process of a buyer (including a potential buyer who may become a customer, i.e., a frequent buyer) in a market. She becomes involved in buyer-buyer communications and the flood of external information. If a peripherally important (mostly authorized) piece of information is found, she becomes concerned with its significance as a chance. This process can be made outside the store by simply chatting about desirable products. Through further communications, still buyer-to-buyer or with salesclerk this time, she becomes concerned with central information i.e., the explanation why a certain product is beneficial for her, and goes closer to the decision to buy it.

In social psychology, it has been known that a simple rumor can be diffused to satisfy people with uncertain decision, which is here reflected to the step at which peripheral information prevails before central information works, i.e., when human is no more confident of the chance chased in the last cycle of interpretation and the concern with the new chance is not fixed yet. The model of diffusion (Rogers, 1962), where people differ in the reaction speed to innovative ideas, also corresponds to the step-wise growth of people's concern in the communications about a new appearance of chance. It has been pointed out that a dialogue, where chance discovery can be made by a group of people, includes the nature of *inquiry* where participants collaborate to find a new solution to a problem (because it is unknown how one can manage a new chance), and of *persuasion* where a participant seeks to persuade other(s) to accept a new proposal (McBurney and Parsons, 2001).

According to the formalization by MacBurney and Parsons, an essential locution of participants in the communication for chance discovery introduces a sharable value-criteria (the recognized relevance of an event to benefit or loss) which makes an opportunity to unify multiple communities.

If the new value-criteria is about products attracting multiple communities, ones leading each community will be unified first because they are ready to touch new products and contact each other to talk about how the communities can contribute to each other for taking the full advantage of the product. This can merge the communities to form a new market sharing the new value criteria, by the diffusion from leaders as illustrated in Fig 1.

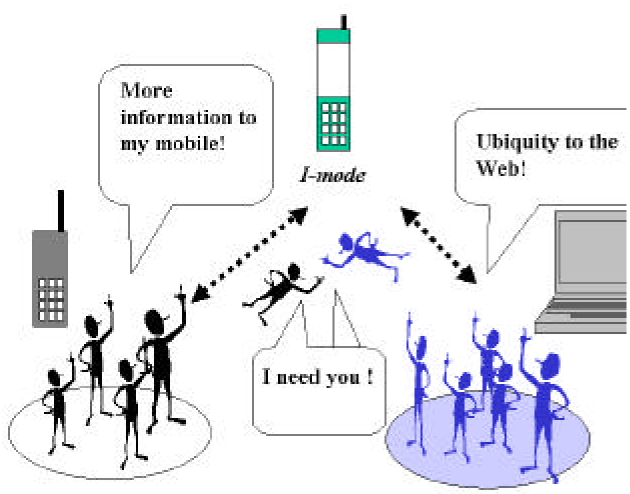

Fig. 3 The diffusion of i-mode, a Japanese mobile phone with interface to the Internet. I-mode appeared as a new value attracting the two communities of cellular- and PC-users, and made them attract each other. This was a cause of the fusion into a big market.

From the analogy between the human-human interaction of ideas and genetic process of evolution, we can regard a new combination of ideas as a breakthrough leading to innovation. For in-depth discussion about this lesson from genetic algorithm and its possible implementation as a real communication (Goldberg, 2002). By building well-defined formalization frameworks, the ability of chance discovery may be transplanted from human to software agents in a long future. However, regarding human as a system learning schema from such logical descriptions (even if some details are missed), we can aim in a nearer future at acquiring the capability of chance discoveries. We are focusing attentions onto communications, as a main engine of chance discovery.

Imagination, Key 2

Human becomes aware of the relevance between one's implicit knowledge (i.e., one knows, but is not conscious of the knowledge) or implicit experience and a confronted situation, using mental imagery (Kosslyn, 1980), i.e., scenes imaginable in conjunction with the confronted situation. It has been shown that talking about scenes where a product is used can be an effective explanation of the utility of the product (Shoji and Hori, 2001). Scenic information has also been proposed as an aid of smooth context shifting. In other words, scenic information shifts human forward in the phase one is weakly concerned with a chance, working as a peripheral clue. In real situation, scenic information and oral communication interplay with each other for creating new aspects of products.

For dealing with the mechanism of finding relevance between the current and memorized experience formally, formulations based on *analogy* seems to be a promising approach. Communications sometimes aid human in taking advantage of analogies among objects in the mind of participants.

Data Mining , Key 3

The most essential reason why existing data mining methods have never realized chance discovery is that both human and its surrounding are complex systems involving unnoticed factors. We can consider at least three types of human vs. environment relations, where:

(1) The environment makes one-way affection to human(s).
(2) Human(s) make one-way affection to his/her/their environment.
(3) Human and the environment interact and affect each other.

For example, let us think of earthquakes occurring from the extremely complex and unknown pressures among land crusts. This system is of type (1), i.e. a pure risk in that we human cannot avoid (by the current state of art in seismology or disaster-prevention technologies) all the risks of the future occurrence of great earthquakes. On the other hand, the same system is of type (2) because human knocks the earth to investigate the mechanism underground, by looking at the response to the shock from the land crust on the surface and the inner part of the earth. Also, in trenching survey human mines the earth for finding 100000-years history of underground events. For doing so effectively, however, human must choose the essential part of the land surface he should act on, considering the past attacks from the land – the chance discovery from this system is of type (3).

For the full understanding of chance discovery as a system, we should make discussions for modeling complex systems where a small event may make a big result in the future. Data mining methods applicable as Key 3 for chance discovery, are based on particular modeling of the real world as a complex system, as some presented in this symposium.

Criteria for Evaluating a Discovered Chance

It is another essential problem how we can measure the significance of a chance. Because a chance might be rare, we can hardly understand its meaning for our course of life. For example, see the communication for chance discovery: First, a few leading people notice a chance in face of its rareness, and an acceptable interpretation is proposed and spread to a number of other people who might share the new value, i.e., the significance of the chance, and the value may grow to be a common value. Thus, the significance of a chance can be measured on *P*, *U*, and *G* below.

Proposability: How reasonable a proposal of decision can be made based on the chance. This can be measured by rate *P* of people who agrees, to a decision proposal based on the chance.

Unnoticability: How difficult the chance is to be noticed without being proposed by other people. This differs from pure *novelty* because an unnoticable chance might have existed without being noticed. This can be measured by *U*, the inverse of the rate of leading people who noticed the significance of the chance, within a short constant time *t* after its first appearance.

Growablility: The speed of increase in the number of people who agrees with the significance of a chance. This can be measured by the growth rate *G*, i.e., (*the number of people who agree with the significance of the chance*) / (*the number of people who noticed the significance of the chance in the beginning time of length t*).

So far, some studies on Chance Discovery have been aiding human awareness of chances of high *P*, *U*, and *G*. For example in (Fukuda and Ohsawa, 2001), relations between foods in breakfast were obtained from the data of meal consumptions by habitants of Tokyo and visualized by *KeyGraph* (a data-mining method summarized later). A group of housewives, exchanging their ideas with being inspired by the graph, enabled the discovery of hidden but influential value criteria of housewives in designing breakfast. In this approach, few discussants' subjective awareness of the significance of rare events spread throughout the group.

This success owes to a revision from the model of knowledge discovery process in Fig.1. A key point was that it came to be considered that a chance can be something unnoticed (rare of new). The model was thus revised to Fig.2, carrying human from ignorance to *concern*, *awareness*, *understanding of the chance*, and *decision/actions*, returning to a new concern with another chance. The group of people stimulated these steps of each other, and data-mining helped externalization of each discussant's ideas. Because the consideration of deepened factors helps in the explanation of the significance of new chances, the process model enabled to discover rare but significant consumptions of foods occurring from deep-level motivations of consumers (Fukuda and Ohsawa, 2001).

What We Did, What We Do

An essential aspect of a chance is that it can be a new seed of significant future changes and actions. The discovery of a rare opportunity may lead to a benefit not experienced before, because people have hardly noticed the benefit and were accustomed to frequent past opportunities. The discovery of a new hazard is indispensable for minimizing the risk, because existing solutions having worked for frequent past hazards may not work this time.

Besides data mining methods for finding rare but important events from time-series, it is also important to draw humans attention to such events and to knowledge for dealing with them, i.e., to make humans ready to catch and manage chances. In this sense, the complex interactions of human-environment, human-human, human-agent (either hardware or software), agent-agent (either hardware or software), and agent-environment are all essential to chance discovery, as well as interactions among various components of the environment. Colleagues of chance discovery have been challenging discoveries of such chances as:

- Signs of great earthquakes in the future, as a complex system where a small cause results in a great effect
- Key-genomes triggering diseases, etc
- Triggers of human, robot, and agent behaviors in dynamic environments

- New products and consumers worth promoting sales
- Signs of critical events in the future
- New answers in questionnaire with high social impacts,
- Words on WWW, attracting hidden public desires,
- Communications for creating new value-criteria
- Leading opinions in on-line communities
- Anomalies with significant impact on the economy
- Visual information aiding our discoveries of chances

etc. In this symposium, presents theories and methods of chance discovery. The new topics include

- Theories on analogy and serendipity as the principle of chance discovery,
- Understanding of changes in the environments
- Managements of human risk-factors as terrorism,
- Theories on belief and confidence
- Theories on and methods for creative communications of human
- Data mining for catching chances from data

These central points of chance discovery will help speakers and the audience in enabling to realize one's decisions in dynamic environments.

References

Goldberg, D.E., 2002. The Design of Innovation: Lessons from and for Competent Genetic Algorithm, Kluwer Academic Publishers Volume 7, Kluwer Series on Genetic Algorithms and Evolutionary Computation.

Nonaka, I. and Takeuchi, H., 1995. The Knowledge Creating Company, Oxford University Press.

Weiss, G.M. and Hirsh,H., 1998. Learning to Predict Rare Events in Event Sequences, in Proc. of KDD-98 pp.359-363.

Turney, P.D., 1995. Cost-sensitive classification: Empirical evaluation of a hybrid genetic decision tree induction algorithm, *Journal of Artificial Intelligence Research*, 2, March, pp..369-409.

Festinger, L., 1957. A Theory of Cognitive Dissonance, Evanson, IL: Row, Peterson.

Nara, Y. and Ohsawa,Y., 2000. Tools for Shifting Human Context into Disasters, in Proceedings of IEEE conference of KES2000.

Petty, R.E., and Caccioppo, J.T., 1981. Attitudes and persuasions: Classic and contemporary approaches, Dubuqute, IA: WC Brown.

Shibata, H. and Hori, K., 2001. A Creativity Support System for Long-term Use in Everyday Life - IdeaManager: Its Concepts and Utilization, New Frontiers in Artificial Intelligence LNAI 2253, pp.455-460, Springer Verlag.

Ohsawa, Y, 2002. Chance Discoveries for Making Decisions in Complex Real World, New Generation Computing, Vol.20 No.2, pp.143-163.

Fayyad, U., Shapiro, G.P. and Smyth, P., 1996. From Data Mining to Knowledge Discovery in Databases, AI magazine, Vol.17, No.3, 37—54.

Allport,G.W and Postman, L., 1947. The Psychology of Rumor, H.Holt.

Rogers, E.M., 1962. Diffusion of Innovations, Free Press

McBurney, P. and Parsons, S., 2001. Chance Discovery Using Dialectical Argumentation, New Frontiers in Artificial Intelligence LNAI 2253, pp.414-424, Springer Verlag.

Kosslyn, S.M., 1980. Image and Mind, Harvard University Press.

Shoji, H. and Hori, K., 2001. Chance Discovery by Creative Communications Observed in Real Shopping Behaviors, New Frontiers in Artificial Intelligence LNAI 2253, pp.462-410, Springer Verlag.

Ohsawa, Y. and Fukuda, H., 2002. Chance Discovery by Stimulated Group of People, Application to Understanding Consumption of Rare Food, Journal of Contingencies and Crisis Management, Blackwell, Vol.10, No.3.