Chance Discovery in Image Diagnosis: - Analysis of Perceptual Cycles -

Junya Morita[1], Kazuhisa Miwa[2], Takayuki Kitasaka[2], Kensaku Mori[2], Yasuhito Suenaga[2], Shingo Iwano[3], Mitsuru Ikeda[4], and Takeo Ishigaki[3]

[1] Graduate School of Human Informatics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan
morita@cog.human.nagoya-u.ac.jp

[2] Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan
miwa@cog.human.nagoya-u.ac.jp, {kitasaka, kensaku, suenaga}@is.nagoya-u.ac.jp

[3] Graduate School of Medicine, Nagoya University
65, Tsurumai-cho, Showa-ku, Nagoya, Aichi, 466-8550, Japan

[4] School of Health Science, Nagoya University
1-20, Daikou-Minami 1, Higashi-ku, Nagoya, Aichi, 461-8673, Japan
a40495@nucc.cc.nagoya-u.ac.jp

Abstract. Image diagnosis is the task in which a physician searches for abnormal findings on the medical images that potentially involve a large amount of anatomical and pathological information, and identifies diseases from such findings. Since this task aims to reduce the risk (e.g., insult or death) that will occur in the near future, it can be seen as a case of chance discoveries. To investigate the cognitive process underlying the task, we conducted an experiment in which two independent physicians diagnosed the same CT images. Analysis of disagreements about final diagnosis and verbal protocols revealed that when observing the CT images on which one disagreed with the other, the physicians engaged in iterative cycles of the image-feature search and the schema construction. The result suggests the importance of investigating cyclic interactions between humans and environments when considering chance discovery.

1 Introduction

A chance is an event that can affect the near future, and that involves opportunities and risks for human decision-making [1]. Discovery of a chance and its impacts on the future are defined as follows:

Meaning of a chance = {X: An encountered event,
Y: A cause of X,
Z: A future scenario caused by Y,
S: A future scenario caused by changing Y.}

Here, focusing on X is called *chance discovery*, and deriving S by searching Y from X is called *scenario creation.* The goals of the studies on chance discovery

Table 1. Materials

	Clinical histories of patients	Aims of investigation
C 1	Pneumonia or previous inflammation in the left lingula.	Screening.
C 2	Shadow in the left upper lobe. Excision of the right upper lobe due to pulmonary tuberculosis. Cancer of lower pharynx.	Examination of the left upper lobe.
C 3	Interstitial pneumonia.	Follow up.
C 4	Lung cancer in the left lung.	Follow up.
C 5	Bladder cancer.	Search for metastasis in the chest.
C 6	Synovial sarcoma of the chest wall.	Search for recurrence or metastasis.
C 7	Pulmonary tuberculosis.	Post-treatment examination
C 8	Emphysema.	Follow up.
C 9	Chronic hepatitis. Mesh shadow in the left lower lung.	Examination of the shadow.
C10	Post-operation for thyroid cancer.	Search for metastasis.

Note. C=Case.

are to find cognitive factors that enlarge the difference from Z to S and to develop tools to support such activities.

In this paper, we report our experimental study on medical image diagnosis as a case of chance discovery. Image diagnosis is a task in which a physician finds abnormal findings on medical images, such as X-ray films or three-dimensional (3-D) X-ray CT (Computed tomography) images, and identifies what disease affects the patient. Based on the above framework, medical image diagnosis is the task that aims to change the future from Z (neglecting diseases) to S (treatment of diseases) by finding X (abnormal findings on the images) and identifying Y (patient's diseases). Therefore, this task can be seen as a typical case of chance discoveries.

Investigating a physician's cognition of image diagnosis is also important from the viewpoint of medical support. Recently, medical images mainly used for diagnosis have been changing from X-ray films to 3-D CT images. Although 3-D information provided by CT images enables a physician to observe abnormal features precisely, such large amounts of information has considerably increased the physician's workload. Since it is important to find subtle signs of malignancy from a large number of images, great expectations have been placed on the development of medical support tools, called *computer aided diagnosis (CAD)* system, to reduce diagnostic cost and improve the accuracy of image diagnosis. Developments so far include, for example, automated detection systems for pulmonary nodules [2–4], quantification systems for emphysema [5], 3-D visualization functions that support perception of anatomical structures in a human body [6–8], and image-matching functions that prevent the missing of changes in follow-up investigations [9–11]. However, most of these comprise "bottom-up" functions that are based on analyzing physical values of images. Therefore the

CAD systems present unnecessary or excessive information to physicians (e.g., false positives in nodule detection).

Thus we consider the investigation into the cognitive factors behind medical image diagnosis to be important for developing more intellectual CAD systems. In investigating a radiologist's cognition, we do not regard it as one-way, bottom-up process that the past CAD systems implicitly assumed, but regard it as an iterative cyclic process that was discussed in Neisser's perceptual cycle theory [12]. Neisser defined the term *perception* as an interaction between an individual and an environment, which is mediated by a *schema*. The term schema, he said, means mental representations for estimating and interpreting environments. A schema is also constructed by both pre-existing experience and available information in environments, and it guides the search for information in environments. That is, perception is dynamically generated through iterative cycles between information searches that are guided by a schema (top-down process) and schema construction based on information in the environments (bottom-up process).

The perceptual cycle theory is suitable for explaining realistic and complex perception influenced by contexts. For example, when one perceives ambiguous figures that potentially involve multiple schemata, pre-existing expectations affect what features will be focused on. Additionally, when one perceives novel objects that one has not been familiar with, a schema will be instantiated incrementally through searching for information about the objects. Thus, we consider that this perceptual cycle theory suitably explains medical image diagnosis, in which physicians view complex and ambiguous medical images.

As mentioned above, medical image diagnosis can be seen as a case of chance discoveries. That is, searching abnormal findings corresponds to the discovery of a chance X, and schema construction corresponds to identify a cause Y. Our goal here is to demonstrate interactions between X and Y in image diagnosis, based on the perceptual cycle theory.

2 Method

2.1 Materials and Procedure

Most of the past experimental studies on image diagnosis have focused mainly on the bottom-up process. Many researchers have conducted experimental studies manipulating physical features of medical images, and the studies have revealed which physical features greatly affect visual searching and final diagnosis [13–15]. Although some researchers have discussed the importance of the top-down process [16], there has been insufficient investigation on interactions between the two processes. Therefore, it is necessary to examine the cognitive process underlying medical image diagnosis.

Our experiment was performed in a room located in the radiological department at Nagoya university, where the participants usually work. In the experiment, the participants were provided with tools that they usually used: a workstation for displaying CT images, a reporting system that defines the format of

Table 2. Examples of coding for medical reports

Physician 1a	Physician 1b	
There is a nodule about 10 mm in diameter in the peripheral right lung.	There is an 1 cm circular nodule in the S3 segment of the right lung.	I
Its boundary is very clear.		I
It also contains an air bronchogram.		I
Endobronchial spread is suspected.	There are satellite lesions surrounding the nodule.	I
	It is highly possible that the nodule is due to inflammation.	S
	There is atelectasis in the lingula.	S
I can see no clear bronchodilation.	*That involves bronchodilation.*	I
Micronodules are detected in the peripheral lingula, the peripheral right lung, and the left upper lobe.	There are many micronodules at the right upper lobe and the S6 of the right lower lung.	I
Tuberculoma is suspected.	Atypical mycobacterial infection is suspected.	S
There are no clear calcifications.		I
The axilla and the mediastinum contain no enlargements of lymph nodes.	There are no abnormalities in the mediastinum.	I

Note. I=IMAGE, S=SCHEMA.

medical reports, and a database system that includes clinical information about the patients, such as electronic charts. To obtain realistic data for the diagnostic process, we selected ten CT images from the ones for which no conclusive diagnosis had been made as task materials (Table 1).

The participants were required to write medical reports about abnormal findings and suspected diseases while viewing the CT images. Our experiment differed from usual diagnosis situations only in instructions for obtaining verbal protocol data by producing a speech sound. The participants were instructed to give talk-aloud protocols and to indicate where they are looking by using a mouse pointer. We recorded all verbalizations, and videotaped the workstation's display.

2.2 Participants

Four physicians who specialize in radiology or pulmonology participated in the experiment. They had at least two years of experience in image diagnosis. Two of the four observed Case 1 to Case 5 individually, and the others observed Case 6 to Case 10 individually. That is, under usual diagnosis situations, each of the cases was diagnosed by two independent physicians. Hereafter, we call the physician who diagnosed Cases 1 to 5 Physician 1a or Physician 1b, and those who diagnosed Cases 6 to 10 Physician 2a or Physician 2b.

Table 3. The number of labels included in the reports

		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Sum
IMAGE	AGREEMENT	6	10	6	12	4	4	2	2	4	2	52
	CORRESPOND	2	0	0	0	0	0	0	0	0	0	2
	NON-CORRESPOND	3	15	6	3	13	0	2	5	4	1	52
SCHEMA	AGREEMENT	2	8	2	4	2	4	2	6	2	2	34
	CORRESPOND	0	4	0	2	2	0	0	0	0	0	8
	NON-CORRESPOND	1	3	3	1	2	0	2	1	3	0	16

Note. C=Case.

Table 4. Agreement rates

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
IMAGE	0.55	0.32	0.51	0.67	0.25	1.00	0.50	0.35	0.53	0.75
SCHEMA	0.75	0.53	0.41	0.58	0.37	1.00	0.66	0.87	0.41	1.00

3 Results and Discussion

3.1 Medical reports

Analysis We analyzed each medical report (it is an official output of diagnosis), dividing the texts of the report into statements according to appearance of predicates. All statements were labeled in the following two ways: First, each statement was labeled as either an "IMAGE" or a "SCHEMA" based on the perceptual cycle theory. The label IMAGE indicates a statement concerning a physical feature of a CT image, such as size of a nodule, density of a shadow, and distribution of shadows. The other label, SCHEMA, indicates a statement concerning a physiological or pathological feature on a CT image, such as identification of the disease, benign/malignant discrimination of a nodule, and speculation about past surgery methods.

The second labeling aims to clarify how and why two physicians disagreed. A disagreement between two physicians suggests that at least one made a false diagnosis, or else one discovered a rare chance that the other missed. In that sense, we consider that investigating disagreements will be important for studies on both chance discovery and CAD. For analyzing disagreements, the statements were divided into three categories. When two physicians gave a consistent value or content in corresponding IMAGE or SCHEMA, the statement was labeled as "AGREEMENT." On the other hand, when two physicians gave a contradicting value or content in a corresponding IMAGE or SCHEMA, the statement was labeled as a "CORRESPONDING DISAGREEMENT." If one physician's statement did not correspond to any statements by the other's, the statement was labeled as a "NON-CORRESPONDING DISAGREEMENT." The label NON-CORRESPONDING DISAGREEMENT indicates that two physicians focused their attention on different image features or schemata. On the other hand, CORRESPONDING DISAGREEMENT indicates clear contradiction between physicians.

Table 2 shows an example of coding. The first and second columns indicate statements of Physician 1a and Physician 1b, respectively, while the third column

Table 5. Corresponding disagreements

	Physician 1a	Physician 1b	
C1	I can see no clear bronchodilatations.	That involves bronchodilatations.	I
C2	A primary cancer is most suspected.	It may be an atypical lung cancer.	S
	There is also the possibility of metastasis from hypopharynx.	There are no metastases from the hypopharynx.	S
C4	Left upper lung was removed.	Left lower lung was removed.	S
C5	Possibility of metastasis from the bladder cannot be denied.	Possibility of metastasis from the bladder is denied.	S

Note. C=Case, I=IMAGE, S=SCHEMA.

indicates IMAGE/SCHEMA labels. Two statements that include a corresponding IMAGE or SCHEMA are aligned with each other. Consistent statements are shown in normal typeface (AGREEMENT) and inconsistent statements in italics (CORRESPONDING DISAGREEMENT). The NON-CORRESPONDING DISAGREEMENT statements are aligned with blank cells.

Results Table 3 shows the number of labels included in the reports of two physicians. In addition, we computed agreement rates of IMAGEs and SCHEMAs in each case (Table 4). The perceptual cycle theory predicted that agreement rates of IMAGEs would positively correlate with the agreement rates of SCHEMAs, because different schemata would lead to different image features, or vice versa. Consistent with this prediction, the rank correlation coefficient between agreement rates of IMAGEs and those of SCHEMAs was marginally significant ($r_s = 0.54, p < .10$).

Next, we investigated how two physicians disagreed. The chi-square test using the number of each label revealed significant differences in AGREEMENT, CORRESPONDING DISAGREEMENT, and NON-CORRESPONDING DISAGREEMENT rates between IMAGE and SCHEMA [$\chi^2(2) = 13.53, p < .01$]. However, the residual analysis did not detect any significant difference in AGREEMENTs between IMAGE and SCHEMA. This result may suggest the necessity of supporting the schema construction as well as the image feature search that the previous CAD systems have supported. The differences in both CORRESPONDING and NON-CORRESPONDING DISAGREEMENT between IMAGE and SCHEMA were significant ($p < .05$), indicating that most of the disagreements of IMAGEs were due to NON-CORRESPONDING DISAGREEMENTs, whereas CORRESPONDING DISAGREEMENTs were relatively frequent in SCHEMA. This result suggests that a method for supporting schema construction must differ from that for searching image features. However, CORRESPONDING DISAGREEMENTs happened only in the cases investigated by Physicians 1a and 1b, therefore, the above discussion must be confirmed in future research.

Finally, we show the contents of CORRESPONDING DISAGREEMENT, which indicate clear contradictions between physicians (Table 5). One contradiction is in the IMAGE of Case 1. The others concern SCHEMA, including a disagreement on surgery methods, which might be due to a lack of information

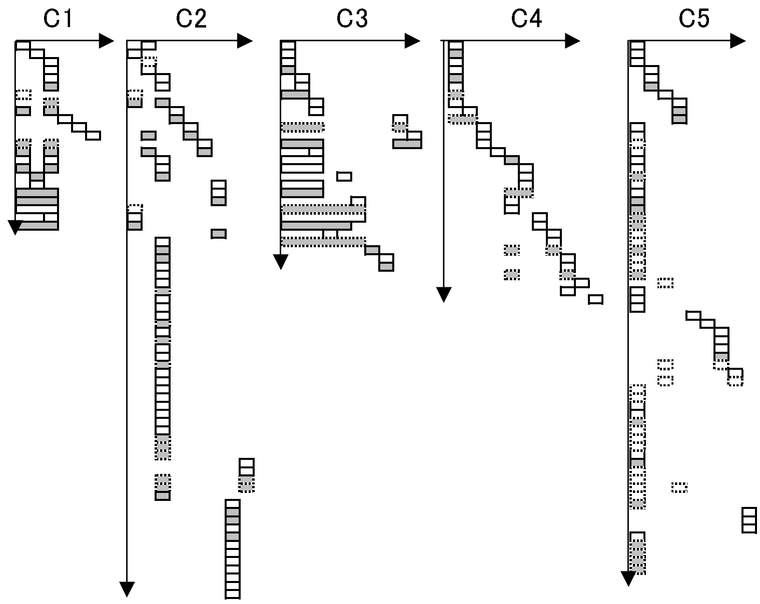

Fig. 1. Diagrams illustrating diagnostic processes.

in clinical history (Case 4). Cases 2 and 5 contained important contradictions on the interpretation of lung metastases in regards to the aims of this investigation (see Table 1).

3.2 Protocol Analysis

What makes their diagnostic results different? To answer this question, using protocol analysis, we have to define the physicians' diagnostic process to their final diagnosis. However, unfortunately, we could not obtain sufficient numbers of verbal protocols in the diagnoses by all of the participants. Perhaps this is because our experiment was performed in a naturalistic experimental setting where participants may have been reluctant to speak talk-aloud protocols. Thus, in the following case study, we analyzed only Physician 1b, who provided a sufficient amount of protocols.

We drew a diagram illustrating the diagnostic process for each case investigated by Physician 1b (Fig. 1). The vertical axes indicate time flow, while the horizontal axes indicate the type of object, such as "a nodule in the right upper lung" or "calcification in the mediastinum." The white cells indicate statements containing image features, and the gray cells indicate statements containing schemata. The cells framed by solid lines indicate newly verbalized image features or schemata, while the cells framed by dotted lines indicate image features or schemata already mentioned. The rectangular cells unified across multiple cells in the same time line indicate a statement synthesizing multiple objects, such as "this nodule is the same kind of the previous nodule."

Table 6. Protocol excerpts from Case 2

If anything, this is an inflammation of lung.	Schema	old
However, I can't deny the possibility of lung cancer.	Schema	old
The inside is very strangely shaped,	Image	old
and opened.	Image	new
If this is a cancer,	Schema	old
the slenderness ratio is too high.	Image	old
An air bronchogram can be seen.	Image	new
Also, there are spiculas.	Schema	new
It may be a lung cancer,	Schema	old
but there are no ground-glass opacities.	Image	new
I suspect... inflammation.	Schema	old
This shadow looks like a satellite lesion.	Image	new

Table 7. Protocol excerpts from Case5

I can see a shadow in S6.	Image	new
High-resolution CT reveals clear scragginess spiculas.	Image	new
there are also ground-glass opacities in the peripheral...(Snip)	Image	new
Lung cancer...	Schema	new
However, it is not a metastasis.	Schema	new
It may be a primary lung cancer or	Schema	new
tuberculosis.	Schema	old
Because it is in S6 where tuberculosis often occurs,	Image	old
I can consider the possibility of tuberculosis.	Schema	old
However, regarding the shape, there are spiculas	Image	old
and the bronchus is disrupted,	Image	old
so I must consider the possibility of lung cancer.	Image	old

Fig. 1 shows that strongly continuous protocols appeared in the second column of Case 2 and the first column of Case 5. Here, the "old" cells increase, meaning that the physician extracted image features (white cells) and constructed schemata (gray cells) iteratively. Detailed protocols are shown in Table 6 (Case 2) and Table 7 (Case 5). In Case 2, based on the two schemata (cancer/inflammation), the physician searched the related features (ground-glass opacity/satellite lesion). In Case 5, using the extracted features (S6/specula/ disrupted bronchi), he evaluated the two schemata (cancer/tuberculosis). The process in Case 2 can regarded as a top-down process because the schemata guided the image-feature search. On the other hand, the process in Case 5 can be considered a bottom-up process because the image features were used to evaluate the schemata.

The above results relate to the results in the final reports. Interestingly, the agreement rates for IMAGE in both Case 2 and Case 5 were lower than those in the other cases (see Table 4). Furthermore, the objects where perceptual cycles were performed repeatedly were identical to the diagnosis on which two physicians disagreed (CORRESPONDING DISAGREEMENT).

4 General Discussion

Analysis of the reports revealed positive correlation between IMAGE and SCHEMA. In the analysis of the verbal protocols we observed iterative cycles of the image-feature search and the schema construction in the cases where disagreements emerged. These results suggest the relation between physicians' disagreements and their iterative cyclic perceptual processes, implying that contradictions in schema lead to differences in extracted image features, or that different image features construct contradicting schemata. That is, the disagreements might increase asymptotically through dynamic iterative cycles between images and schemata. However, we cannot confirm this hypothesis because we were not able to analyze the other's protocols.

The most important implication of our study is that the cases of the iterative cycles on image diagnosis were shown. Since our experimental set-up was almost identical to that of a practical diagnosis situation, our results provide ecologically valid evidence for the perceptual cycle theory. Findings of the cyclic process observed in our study will be also important for studies on chance discovery. As mentioned at the beginning, extracting features from a CT image correspond to the discovery of an event X, and constructing a schema corresponds to the reasoning of a cause Y. Thus, our results suggest that a future scenario S is created through iterative cyclic interactions between X and Y.

Such discussion is consistent with recent studies on chance discovery. For example, Suwa et al. conducted a case study on professional designer's cognitive processes, revealing interactions between the perception of new elements in sketches and the generation of novel ideas [17]. Ohsawa et al. also proposed the double helical model, which aims to explain the interactive process between available data and individuals' hypotheses [18]. An important feature of the double helical model is in proposing a method to support the discovery of chances.They developed a "key graph" that visualizes rare correlations among multiple elements in the environment, and demonstrated that applying the key graph leads to both discovery of a rare chance and elaboration of a hypothesis. Our results are consistent with the above studies, showing a case of interactive processes between data and individuals. Consequently, it should be useful to develop tools that let physicians notice rare features in CT images, as shown by a key graph. Our future challenge is to develop an effective method that supports the perceptual cycles of image diagnosis.

Acknowledgments

The authors thank Dr. T. Nakamura and Dr. H. Fukushima of Nagoya University Hospital, Dr. N. Inoue of Tosei General Hospital for participating in the diagnosis experiment. They also thank the members of our laboratory at Nagoya University for their collaboration. This work was supported in part by the Grant-In-Aid for 21st COE program provided by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government.

References

1. Ohsawa, Y.: Chance discoveries for making decisions in complex real world. New Generation Computing **20** (2002) 43–163
2. Kanazawa, K., Kawata, Y., Niki, N., et al.: Computer-aided diagnosis for pulmonary nodules based on helical CT images. Computerized Medical Imaging & Graphics, **22**, (1998) 157–167
3. Yamamoto, S., Jiang, H., Matsumoto, M., et al.: Image Processing for Computer-Aided Diagnosis of Lung Cancer by CT (LSCT). Proceedings of the Third IEEE Workshop on Applications of Computer VIsion, (1996) 236–241
4. Giger, M. L., Bae, K. T., and MacMahon, H.: Computerized detection of pulmonary nodules in computed tomography images. Invest Radiology. **29** (1994) 459–465
5. Mori, K., Aiguchi, T., Ikezaki, M., et al.: CAD system for quantitative evaluation of chronic obstructive pulmonary disease based on 3-D CT images. Proceedings of the 17th International Congress and Exhibition on Computer Assisted Radiology and Surgery (CARS), (2003) 1049–1054
6. Vining, D. J., Thitrin, R. Y., Haponik, E. F., et al.: Virtual Bronchoscopy. Radiology, 193(P), Supplement to Radiology (RNSA Scientific Program). 1994
7. Mori, K., Hasegawa, J., Toriwaki, J., et al.: Automated Extraction and Visualization of Bronchus from 3D CT Images of Lung. Proc. of 1st International Conference on Computer Vision, Virtual Reality and Robotics in Medicine, (1995) 542–548
8. Roggala, R., Scheltinga, J. T., Hamm, B., eds.: Virtual endoscopy and related 3D techniques. Springer, Berlin (2001)
9. Maurer, C. R., Aboutanos, B., Dawant, B. M., et al.: Registration of 3-D Image Using Weighted Geometrical Features. IEEE Trans. on Medical Imaging, **15**, 6, (1996) 827–849
10. Rueckert, D., Sonoda, L. I., Hayes, C., et al.: Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. IEEE Trans. on Medical Imaging, **18**, 8, (1999) 712–721
11. Sato, H., Ukai, Y., Niki, N., et al.: Computer-Aided Diagnosis System for Comparative Reading of Helical CT Images for the Detection of Lung Canver. IEICE Trans. Inf. & Syst. **Vol.E84-D**, 1, (2001) 161–170
12. Neisser, U. : Cognition and reality. (1978) W.H.Freeman and Company.
13. Matsumoto, T., Fukuda, N., Tsuchikawa, M., et al.: Observer Performance Study for CT-Image Reading of One Slice or Multislice by the Cine Display Mode of CRT System -An Application of the Diagnostic-Dynamic Characteristic (DDC) Model. Proceedings of the 17th International Congress and Exhibition on Computer Assisted Radiology and Surgery (CARS), (2001) 1190
14. Krupinski, E. A., Radvany, M, Levy, A., et al.: Enhanced Visualization Processing: effect on workflow. Academic Radiology, 8, (2001) 1127–1133.
15. Krupinski, E. A. and Roehrig, H.: Pulmonary nodule detection and visual search: P45 and P104 monochrome versus color monitor displays. Academic Radiology, 9 (2002) 638-645.
16. Krupinski, E. A.: The future of image perception in radiology: synergy between humans and computers (Guest Editorial). Academic Radiology, **10**, 1, (2003) 1–3.
17. Suwa, M., Gero, J., and Purcell, T.: Unexpected discoveries and S–invention of design requirements: important vehicles for a design process Design studies **21** (2000) 539–567
18. Ohsawa, Y. and Nara, Y.: Decision process modeling across internet and real world by double helical model of chance discovery. New Generation Computing **21** (2003) 109–121