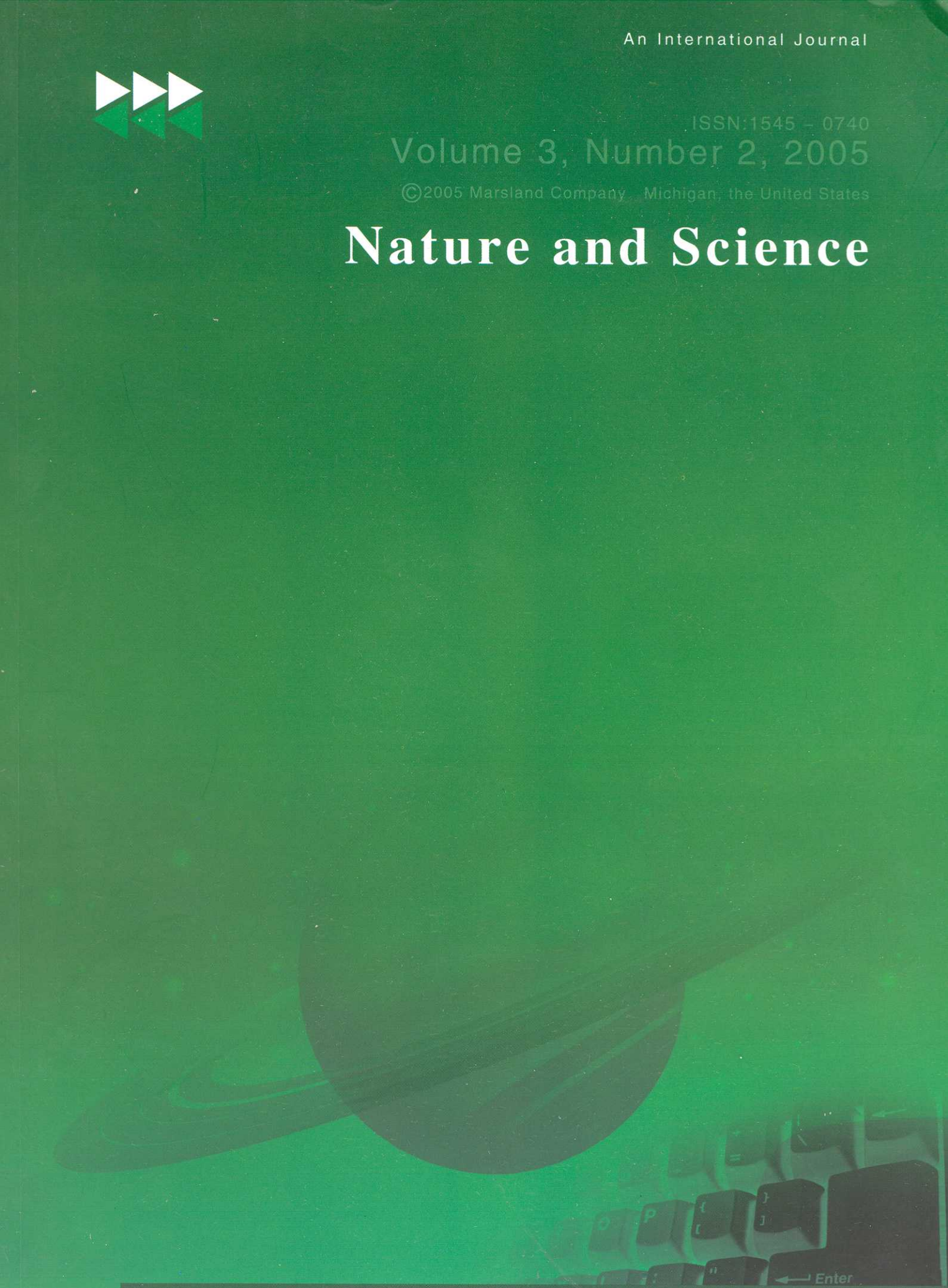
An International Journal

ISSN:1545 - 0740
Volume 3, Number 2, 2005

Nature and Science



Nature and Science

The **Nature and Science** is an international journal with a purpose to enhance our natural and scientific knowledge dissemination in the world under the free publication principle. Any valuable papers that describe natural phenomena and existence or any reports that convey scientific research and pursuit are welcome, including both natural and social sciences. Papers submitted could be reviews, objective descriptions, research reports, opinions/debates, news, letters, and other types of writings that are nature and science related. The journal is calling for papers and seeking co-operators and editors as well.

Introductions to Authors

1. General Information

(1) Goals: As an international journal published both in print and on internet, Nature and Science is dedicated to the dissemination of fundamental knowledge in all areas of nature and science. The main purpose of Nature and Science is to enhance our knowledge spreading in the world under the free publication principle. It publishes full-length papers (original contributions), reviews, rapid communications, and any debates and opinions in all the fields of nature and science.

(2) What to Do: The Nature and Science provides a place for discussion of scientific news, research, theory, philosophy, profession and technology - that will drive scientific progress. Research reports and regular manuscripts that contain new and significant information of general interest are welcome.

(3) Who: All people are welcome to submit manuscripts in any fields of nature and science.

(4) Publication Costs: US$30 per printed page of an article to defray costs of the publication will be paid by the authors when the submission or after the acceptance. Extra expense for color reproduction of figures will be paid by authors (estimate of cost will be provided by the publisher for the author's approval).

(5) Journal Copies to Authors: One hard copy of the journal will be provided free of charge for each author.

(6) Additional Copies Bought by Authors: Additional hard copies could be purchased with the price of US$4/issue.

(7) Distributions: Web version of the journal is freely opened to the world without any payment or registration. The journal will be distributed to the selected libraries and institutions for free. US$5/issue hard copy is charged for the subscription of other readers.

(8) Advertisements: The price will be calculated as US$400/page, i.e. US$200/a half page, US$100/a quarter page, etc. Any size of the advertisement is welcome.

2. Manuscripts Submission

(1) Submission Methods: Electronic submission through email is encouraged and hard copies plus an IBM formatted computer diskette would also be accepted.

(2) Software: The Microsoft Word file will be preferred.

(3) Font: Normal, Times New Roman, 10 pt, single space.

(4) Indent: Type 4 spaces in the beginning of each new paragraph.

(5) Manuscript: Don't use "Footnote" or "Header and Footer".

(6) Cover Page: Put detail information of authors and a short title in the cover page.

(7) Title: Use Title Case in the title and subtitles, e.g. "**Debt and Agency Costs".**

(8) Figures and Tables: Use full word of figure and table, e.g. **"Figure 1.** Annul Income of Different Groups**", Table 1. Annual Increase of Investment".**

(9) References: Cite references by "last name, year", e.g. "(Smith, 2003)". References should include all the authors' last names and initials, title, journal, year, volume, issue, and pages etc.

Reference Examples:

Journal Article: Hacker J, Hentschel U, Dobrindt U. Prokaryotic chromosomes and disease. Science 2003;301(34):790-3.

Book: Berkowitz BA, Katzung BG. Basic and clinical evaluation of new drugs. In: Katzung BG, ed. Basic and clinical pharmacology. Appleton & Lange Publisher. Norwalk, Connecticut, USA. 1995:60-9.

(10) Submission Address: editor@sciencepub.net, Marsland Company, P.O. Box 21126, Lansing, Michigan 48909, The United States, 517-980-4106.

(11) Reviewers: Authors are encouraged to suggest 2-8 competent reviewers with their name and email.

2. Manuscript Preparation

Each manuscript is suggested to include the following components but authors can do their own ways:

(1) Title page: including the complete article title; each author's full name; institution(s) with which each author is affiliated, with city, state/province, zip code, and country; and the name, complete mailing address, telephone number, facsimile number (if available), and e-mail address for all correspondence.

(2) Abstract: including Background, Materials and Methods, Results, and Discussions.

(3) Key Words.

(4) Introduction.

(**5) Materials and Methods**.

(**6) Results**.

(**7) Discussions**.

(**8) References**.

(**9) Acknowledgments**.

Journal Address:

Marsland Company
P.O. Box 21126
Lansing, Michigan 48909,
The United States
Telephone:(517) 980-4106
E-mail: editor@sciencepub.net
Homepage: http://www.sciencepub.org

Nature and Science

(Quarterly, Started in 2003)

Volume 3 - Number 2 (Cumulated No. 7), July 1, 2005

CONTENTS

Apoptosis

Hongbao Ma *, Kuan-Jiunn Shieh **, George Chen *

* Michigan State University, East Lansing, MI 48824, USA,
Telephone: 517-432-0623; Email: hongbao@msu.edu

** Department of Chemistry, Chinese Military Academy, Fengshan, Kaohsiung, Taiwan 830, ROC.
Telephone: 011-886-7742-9442; Email: chemistry0220@gmail.com

Abstract: There are two ways for cell dying: (1) By injury or disease. (2) Suicide. Programmed cell death is also called apoptosis - cell suicide. Apoptosis is that the cells undergo death to control cell proliferation. There are 3 different mechanisms by which a cell commits suicide by apoptosis: (1) Generated by signals arising within the cell; (2) Triggered by death activators binding to receptors at the cell surface; (3) Triggered by disadvantaged environment. Besides animal, plant also performs the apoptosis. [Nature and Science. 2005;3(2):1-4].

Key Words: apoptosis; cell; death

Contents

Introduction

For all the things existed, including the life cells in the earth and universe itself, there is a time to live and a time to die. There are two ways in which cells die: (1) Killed by injury or disease. (2) Suicide. Programmed cell death is also called apoptosis, which is cell suicide. Apoptosis is a mechanism by which cells undergo death to control cell proliferation or in response to DNA damage. Some types of cancer cells, such as B-cell chronic lymphocytic leukemia, follicular lymphoma (Tsujimoto, 1985) and tumors infected by human T-cell

leukemia/lymphoma virus-1 (Hengartner, 2000) are characterized by defects in apoptosis leading to immortal clones of cells. Other malignancies have defects in the apoptotic regulatory pathways such as p53 (Kaufmann, 2001). Apoptosis plays an important role in the pathogenesis of cholesteatoma (Olszewska, 2005). The understanding of apoptosis has provided the basis for novel targeted therapies that can induce death in cancer cells or sensitize them to established cytotoxic agents and radiation therapy (Ghobrial, 2005).

Kiseleva, Krylov, Lyudyno and Suvorov recently showed that Vascular endothelial growth factor modulated mitogen-induced proliferation of thymocytes and stimulated spontaneous apoptosis in intact thymus cells (Kiseleva, 2005).

Stoka and Turk described an emphasis to the newly discovered role of lysosomal cathepsins in apoptotic pathways (Stoka, 2005).

1 Death by Injury

Life cells can be killed by injury, such as mechanical injury and toxic environment injury. In the cell dying procedure by injury, the cells undergo a characteristic series of changes, such as the three changes: (1) Cells swell because the disability of the plasma membrane controlling the passage of ions and water. (2) Leak out of cell contents. (3) Inflammation of surrounding tissues. For example, various factors, including nucleotide depletion, electrolyte imbalance, reactive oxygen species, endonucleases, disruption of mitochondrial integrity, and activation of various components of the apoptotic machinery, have been implicated in renal cell vulnerability (Padanilam, 2003).

2 Death by Suicide – Apoptosis

Cell death by suicide is called apoptosis. The apoptosis has the following characterizations: (1) Cell movement is disabled. (2) Nuclear acids and degrade. (3) Cells shrink. (4) Cells appear non-regular shapes on their surface. (5) Water is lost. (6) Cell organelle break down. (7) Cells break into small fragments. Then the cells suicide and died.

3 Why Should a Cell Need Apoptosis as the Suicide?

3.1 There could be two reasons for suicide

(1) Mitosis needs apoptosis. For example: A. The frog re-absorbs the tadpole tail when its metamorphosis changes into a frog. B. The formation of fingers and toes requires the removal of the tissue by apoptosis. C. The sloughing off of the inner lining of the uterus by apoptosis. D. The formation of proper connections between neurons in brain requires surplus cells to be eliminated by apoptosis. (2) Apoptosis is needed to remove cells that hurt the integrity of the organism. For example: Cells infected with bacteria, viruses or other toxic organisms.

3.2 DNA damage causes apoptosis

Genome damage can cause a cell disrupted properly. This is an auto-regulation process that is important for the cell development.

3.3 Cell immune system causes apoptosis

In the immune procedure, the apoptosis plays an important role. To remove non-needed parts of the cells by apoptosis is happened in programmed cell death procedure.

3.4 Cancer cells

For cancer therapy, radiation and chemicals could induce apoptosis. Tipping the balance between cell death and proliferation in cell survival could cause tumor formation. And, killing cancer cells by chemical therapies depends on activation of apoptosis procedures. Therefore, failure to undergo apoptosis in the therapy could result in cancer resistance (Fulda, 2004).

4 What Makes the Apoptosis?

4.1 The balance between positive and negative signals

The balance between the positive signals and negative signals plays important roles in the apoptosis procedures. This is important in the cell's life.

4.2 Positive signals

For the life, cells need to receive continuous stimulation from other cells and, for many, continued adhesion to the surface on which they are growing. These stimulation signals are called signals, such as: growth factors, interleukin-2, etc.

4.3 Negative signals

When the apoptosis happens, the following results could occur: (1) Oxidant level within the cell is increased. (2) DNA is damaged by these oxidants or other agents like ultraviolet light, x-rays, etc. (3) non-properly folded proteins are accumulated. (4) Molecules that bind to specific receptors on the cell surface appear. (5) Some signals related to apoptosis begin to appear.

5 The Mechanisms of Apoptosis

There are 3 different mechanisms of apoptosis: (1) Generated by signals in a cell; (2) Triggered by death factors on cell surface. (3) Triggered by toxic factors.

5.1 Apoptosis caused by internal signals

These signals could be Bcl-2, Apaf-1 (apoptotic protease activating factor-1), Bax, cytochrome c, caspase 9, ATP, etc.

5.2 Apoptosis caused by external signals

Fas, FasL, TNF, TNF receptor, etc could be related to the by external signals.

6 Apoptosis and Cancer

There is closed relationship between apoptosis and cancer and many studied have been reported for this relationship. The research results show that several human papilloma viruses could cause cervical cancer. One of them produces a protein (E6) that binds and inactivates the apoptosis promoter p53. These actions make the cell more resistant to apoptosis.

6.1 Some B-cell leukemias and lymphomas block apoptotic signals.

6.2 Melanoma cells avoid apoptosis by inhibiting the expression of the gene encoding Apaf-1.

6.3 Some cancer cells, secrete elevated levels of a soluble "decoy" molecule that binds to FasL, plugging it up so it cannot bind Fas.

6.4 Some cancer cells express high levels of FasL and they can kill cytotoxic T cell that kills them because cytotoxic T cells also express Fas.

7 Apoptosis and AIDS

Infection with the human immunodeficiency virus type 1 (HIV-1) leads to progressive immunodeficiency. The loss of immune competence is associated with declines in both the functionality and the number of CD4+ lymphocytes. In vitro explorations on the cytopathic effects of HIV-1 have yielded a wealth of potential triggering events, and signaling and effector pathways leading to apoptosis. The types of pro- and anti-apoptotic stimuli that have been associated with HIV-1 are multiple and often appear overlapping or even contradictory (Roshal, 2001). When the infected T cell encounters an uninfected one, the interaction of FasL with Fas on the uninfected cell kills it by apoptosis.

8 Apoptosis in Plants

Besides animal, plant also performs the apoptosis, that turns on a system of programmed cell death. The mechanism differs from that in animals although it, too, involves a protease that — like caspases — cleaves other proteins at Asp and Asn residues. Activation of this enzyme destroys the central vacuole, which is followed by disintegration of the rest of the cell. ()

9. Apoptosis and Human Society

Are there apoptosis in the human social system? There must be yes. All the social systems should have apoptosis mechanisms to program their life. No social system can survive forever, no matter it is a socialism society or a capitalism society, and no matter it is the current Chinese social system or the current American social system. The apoptosis principle must applicable to the human society.

10. Discussions

No matter it is a unique or a universal phenomenon of the earth life in the universe, all the earth life should be died, although we can imagine that the single cell life body can live “forever” through the forever dividing, I personal consider that all the life must die. Even the universe could be die. Are there programs to control this birth and die? Nobody knows. The studies of apoptosis are not only a bioscience application, but also a scientific conception and a philosophy. This concept of apoptosis reveals that the life has a law to make the life to automatic dying.

Apoptosis is the biochemical protocol for the life body in the earth. If a life is come out in the world, it automatically creates a way that the life tries to survive as longer as possible, which likes that we human wish.

But, everybody will be old and died finally, including all the human men/women.

Apoptosis is a universal event in the universe, that happens in all the life bodies and azoic things in the universe, including the universe itself. To understand apoptosis clearly will be important to the health science and also meaningful to the understand of the basic nature laws.

Is there any possibility for us to find some ways to remove apoptosis, so that we can live forever? From the theory point, everything is possible. It is possible for us to find a way to completely remove the apoptosis of human bodies. When this goal is accomplished, people can live at a specific age forever. Interesting? Exciting? The answer is yes, of course!!!

Correspondence to:
Hongbao Ma
B410 Clinical Center
Michigan State University
East Lansing, MI 48824, USA
Telephone: (517) 432-0623
Email: hongbao@msu.edu

References

[1] Fulda S, Debatin KM. Apoptosis signaling in tumor therapy. Ann N Y Acad Sci 2004;1028:150-6.
[2] Ghobrial IM, Witzig TE, Adjei AA. Targeting apoptosis pathways in cancer therapy. CA Cancer J Clin 2005;55:178-94.
[3] Hengartner MO. The biochemistry of apoptosis. Nature 2000;407:770–6.
[4] Kaufmann SH, Hengartner MO. Programmed cell death: alive and well in the new millennium. Trends Cell Biol 2001;11:526–34.
[5] Kiseleva EP, Krylov AV, Lyudyno VI, Suvorov AN. Effect of VEGF on mouse thymocyte proliferation and apoptosis in vitro. Bull Exp Biol Med 2005;139(5):576-9.
[6] Olszewska E, Chodynicki S, Chyczewski L. Apoptosis in the pathogenesis of cholesteatoma in adults. Eur Arch Otorhinolaryngol 2005.
[7] Padanilam BJ. Cell death induced by acute renal injury: a perspective on the contributions of apoptosis and necrosis. Am J Physiol Renal Physiol 2003;284(4):F608-27.
[8] Roshal M, Zhu Y, Planelles V. Apoptosis in AIDS. Apoptosis 2001;6(1-2):103-16.
[9] Stoka V, Turk B, Turk V. Lysosomal cysteine proteases: structural features and their role in apoptosis. IUBMB Life. 2005;57(4-5):347-53.
[10] Tsujimoto Y, Cossman J, Jaffe E, Croce CM. Involvement of the bcl-2 gene in human follicular lymphoma. Science 1985;228:1440–3.

Physics and Cosmology Based on Absolute Motion

Ken H. Seto

KHS Publishing, 260 Yorkshire Lane, Xenia, OH 45385, USA; kenseto@erinet.com;
http://www.erinet.com/kenseto/book.html; Telephone: 937-372-5298

Abstract: A new model of the universe called Model Mechanics has been formulated. Model Mechanics explains all the forces of nature with the same mechanism and thus it is able to unite all the forces of nature naturally. Model Mechanics enables us to describe all the processes and interactions in terms of absolute motion of S-Particle or S-Particle systems (a new description of matter) in the E-Matrix (a new description of physical space). Gravity is the attractive force between two objects having the same direction of absolute motion in combination with the repulsive CRE force creates by the same absolute motion of the objects. This explains why the force of gravity is so weak compared to the other forces. [Nature and Science. 2005;3(2):5-20].

Key Words: Physics and cosmology； Absolute motion

Introduction

A new model of the universe called Model Mechanics has been formulated. Model Mechanics explains all the forces of nature with the same mechanism and thus it is able to unite all the forces of nature naturally. In cosmology, Model Mechanics provides solutions to the following problematic cosmological observations: the observed accelerated expansion of the far reached regions of the universe; the observed rotational curves of galaxies disagree with the predictions of GRT; the observed paths of travel of the space crafts Pioneer 10 and 11 disagree with the predictions of GRT and the observable universe appear to have a much larger horizon than it is allowed by its observed age.

Model Mechanics leads to a new theory of gravity called Doppler Theory of Gravity (DTG) and unites gravity with the electromagnetic and nuclear forces naturally [1,2]. It also leads to a complete theory of motion called IRT (Improved Relativity Theory). IRT includes SRT as a subset. However, unlike SRT, the equations of IRT are valid in all environments, including gravity.

Model Mechanics is based on the existence of absolute motions of objects in a stationary and structured light-conducting medium called the E-Matrix. New interpretations of past experiments such as the Michelson-Morley experiment (MMX) [3] enables us to conclude that on earth the direction of absolute motion is in the vertical direction. Based on that interpretation, proposed new experiments to detect absolute motions in the E-Matrix have been formulated.

Model Mechanics Description of The Current Universe

Model Mechanics supposes that a stationary substance, called the 'E-Matrix', occupies all of pure-space (void) in our Universe. Subsequently, we perceive the E-Matrix as space. The E-Matrix, in turn, is composed of 'E-Strings', which are very thin three-dimensional elastic objects, of diameter estimated at 10^{-33} cm. The length of an E-String is not defined. Away from matter, the E-Strings are oriented randomly in all directions. This means that a slice of the E-Matrix in any direction will look the same. Near matter, the E-Strings are more organized: some emanate from the matter, and the number of these passing through a unit area followed the well-known inverse square law of physics. The E-Strings repel each other. This means that there is an unknown outside force that is compacting them together. The repulsive force and the compacting force are in equilibrium. This state of the E-Matrix allows massive matter particles to move freely within it. The motion of a matter particle or particle system in the E-Matrix is called 'absolute motion'. The absolute motion of matter in the E-Matrix will distort the local E-Strings. The E-Strings will recover to the

non-distorted state after the passage of the matter particles. Light consists of wave-packets in neighboring E-Strings. On its way toward its target, a wave-packet will follow the geometry of these neighboring E-Strings. This description of light embodies 'duality', *i.e.* light possessing properties of a mass-bearing particle as well as a wave packet.

With this description of the E-Matrix (space), the next relevant question is: What is matter? All stable and visible matter is made from three basic particles: the electrons, the up quarks, and the down quarks. The protons and neutrons in the nuclei of all the atoms are made from the up quarks and the down quarks. The electrons orbit around the nuclei to complete the picture of all the atoms. The three basic particles are, in turn, made from one truly fundamental mass-bearing particle, called the 'S-Particle'. An S-Particle is a three-dimensional spherical object. It is repulsive to the E-Strings surrounding it and therefore its motion in the E-Matrix is maintained. An S-Particle orbiting around an E-String in the helical counterclockwise direction is an electron. This motion of the S-Particle is the fastest in the E-Matrix, and it gives rise to one unit of negative electric charge. A down quark is also an S-Particle orbiting around an E-String in the helical counterclockwise direction. The speed of its orbiting motion is only 1/3 that of the electron, giving the down quark a negative 1/3 electric charge. An up quark is an S-Particle orbiting around an E-String in the helical clockwise direction at 2/3 the speed of the electron, resulting a 2/3 positive electric charge.

There is one more stable basic particle: the electron neutrino. An electron neutrino has no detectable electric charge, and therefore it does not interact with the other three charged basic particles. It is composed of an S-Particle orbiting around an E-String in the counterclockwise direction like the electron. However, it is moving in a corkscrew like motion away from the charged basic particles. This means that the distortion in the E-Matrix created by the absolute motion of the S-Particle of the electron neutrino will have already dissipated by the time the charged basic particles are ready to interact with it. This is the reason why the electron neutrino does not interact electromagnetically with the charged basic particles.

This simple description of all stable visible matter can answer the thorny question: What *is* the mass of a basic particle? The answer is: mass is the evidence of the orbiting diameter of its S-Particle. Those S-Particles that are not in a state of orbiting motion do not possess any electric charge and therefore they will not interact with the basic charged particles electrically. They will, however, interact with them gravitationally. They are the dark matters predicted by the astronomers.

The next relevant question is: what are the processes that give rise to all the forces between matter particles? The proposed answers to this question are as follows:

1) All the processes of Nature are the result of matter particles reacting to the geometries of the E-Strings (*i.e.* distortions or waves) to which they are confined because of their orbiting motions around these E-Strings.
2) Absolute motions of two objects in the same direction in the E-Matrix will cause the objects to converge to each other--an attractive force. Absolute motions of two objects in the opposite directions in the E-Matrix will cause the objects to diverge from each other--a repulsive force.

This completes the Model Mechanical description of our current universe. All the particles, all the forces and all the processes of nature can be derived from this one description. Model Mechanics replaces the math constructs of space-time and field/virtual particle with the E-Matrix and the distortions or waves in the E-Matrix. The math of the Standard Model is compatible with Model Mechanics and therefore we can use it in combination with the Model Mechanical interpretations to give us better explanations for all the processes of nature.

IRT: Improved Relativity Theory

Special Relativity Theory (SRT) posits that the speed of light is a universal constant in all inertial frames, but suppose the speed of light is not a universal physical constant as asserted by the SRT, but rather a constant mathematical ratio as follows:

$$\frac{\text{light path length of rod}\left(299,792,458\,\text{m}\right)}{\text{absolute time content of clock second co-moving with rod}}$$

This new interpretation for the speed of light revives the discarded notion of absolute time and physical space. It also makes the notion of absolute time and space compatible with SRT. Based on this interpretation for light speed, a new theory has been formulated for motion: Improved Relativity Theory (IRT). IRT includes SRT as a subset, but its equations

are valid in all environments—including gravity. The following is a description of IRT:

The Postulates of IRT:

1. The laws of physics based on a clock second and a light-second to measure length are the same for all observers in all inertial reference frames.
2. The speed of light in free space based on a clock second and a light-second to measure length has the same mathematical ratio c in all directions and all inertial frames.
3. The laws of physics based on a defined absolute second and the physical length of a rod is different in different frames of reference.
4. The one-way speed of light in free space based on a defined absolute second and the physical length of a measuring rod has a different mathematical ratio for light speed in different inertial frames. The speed of light based on a defined absolute second and the physical length of a measuring rod is a maximum in the rest frame of the E-Matrix.

The Consequences of these Postulates:

1. The speed of light is not a universal constant. It is a constant math ratio as follows:
   Light path length of rod (299,792,458 m)/the absolute time content for a clock second co-moving with the rod.
   The detailed explanation of this new definition:
   By definition the speed of light in the rest frame of the E-Matrix is as follows:
   Light path length of rod in the E-Matrix frame = 299,792,458m.
   The absolute time content for a clock second in the E-Matrix frame = 1 E-Matrix frame clock second.
   Therefore the speed of light in the E-Matrix frame is: 299,792,458m/1 E-Matrix clock second
   The speed of light in any frame moving in the stationary E-Matrix is determined as follows:
   The light path length of rod in the moving frame= γ (299,792,458m)
   The absolute time content for a moving clock second= γ (E-Matrix clock seconds)
   Therefore the speed of light in any moving frame in the stationary E-Matrix is as follows:
   γ (299,792,458m) / γ (E-Matrix clock seconds).
   This is reduced to a constant math ratio of: 299,792,458m/1 E-Matrix clock second
2. The physical length of a rod remains the same in all frames of reference. The light path length of a rod changes with the state of absolute motion of the rod. The higher is the state of absolute motion the longer is its light path length.
3. The rate of a clock is dependent on the state of absolute motion of the clock. The higher is the state of absolute motion the slower is its clock rate.
4. Absolute time exists. The relationship between clock time and absolute time is as follows: A clock second will contain a different amount of absolute time in different states of absolute motion (different frames of reference). The higher is the state of absolute motion of the clock the higher is the absolute time content for a clock second.
5. Simultaneity is absolute. If two events are simultaneous in one frame, identical events will also be simultaneous in different frames. However the time interval for the simultaneity to occur will be different in different frames. This is due to that different frames are in different states of absolute motion.
6. Relative motion between two observers A and B is the vector difference of the vector component of A's absolute motion and the vector component of B's absolute motion along the line joining A and B.

The Math of IRT:

1) The time dilation (contraction) or expansion equations:

A and B are in relative motion from observer A's point of view:

$$T_{ab} = T_{aa}\left(\frac{F_{aa}}{F_{ab}}\right) \qquad (1)$$

OR

$$T_{ab} = T_{aa}\left(\frac{F_{ab}}{F_{aa}}\right) \qquad (2)$$

T_{aa} = A clock time interval in observer A's frame as measured by A

T_{ab} = A's prediction of B's clock time interval for an interval of T_{aa} in his frame.

F_{aa} = Frequency of a standard light source in A's frame as measured by A.

F_{ab} = Frequency of an identical light source in B's frame as measured by A. If F_{ab} is not constant the mean value is used.

Note: Even though T_{aa} and T_{ab} are two different clock time intervals but both of these clock time intervals contain the same amount of absolute time.

2)The light path length contraction or expansion equations:

$$L_{ab} = L_{aa}\left(\frac{F_{aa}}{F_{ab}}\right) \qquad (3)$$

OR

$$L_{ab} = L_{aa}\left(\frac{F_{ab}}{F_{aa}}\right) \qquad (4)$$

L_{aa} = The light path length of a rod in A's frame as measured by A.

L_{ab} = The light path length of an identical rod in B's frame as predicted by A.

Note: Even though L_{aa} and L_{ab} are two different light path lengths but these two light path lengths are derived from identical rods that have the same physical rod lengths. The different light path lengths are the results of different states absolute motion of the rods.

3)**The Coordinate Transformation Equations:**

$$x' = \frac{f_{aa}}{f_{ab}}\left[x + t\left(f_{aa} - f_{ab}\right)\lambda\right] \qquad (5)$$

$$t' = \frac{f_{aa}}{f_{ab}}\left[t + x\left(\frac{\left(f_{aa} - fab\right)}{\lambda f_{aa}^2}\right)\right] \qquad (6)$$

OR

$$x' = \frac{f_{ab}}{f_{aa}}\left[x - t\left(f_{aa} - f_{ab}\right)\lambda\right] \qquad (7)$$

$$t' = \frac{f_{ab}}{f_{aa}}\left[t - x\left(\frac{\left(f_{aa} - f_{ab}\right)}{\lambda f_{aa}^2}\right)\right] \qquad (8)$$

A is the observer's frame (unprimed) and B is the observed frame (primed).

f_{aa} = The instantaneous frequency measurement of a standard light source in A's frame as measured by A.

f_{ab} = The instantaneous frequency measurement of an identical light source in B's frame as measured by A.

λ = The wave length of the standard light source in A's frame as measured by A.

These coordinate transform equations are valid in all environments--including gravity. This means that IRT will give matching predictions as GRT and at the same time includes SRT as a subset.

4)**Momentum of an object**:

$$p = M_o\lambda\left(F_{aa} - F_{ab}\right) \qquad (9)$$

5)**Kinetic Energy of an object:**

$$K = M_o\lambda^2 F_{aa}^2\left(\frac{F_{aa}}{F_{ab}} - 1\right) \qquad (10)$$

6)**Energy of a single particle:**

$$E = M_o\lambda^2 F_{aa}^2 \qquad (11)$$

7)**Gravitational Red or Blue Shift:**

$$\Delta F_{aa} = F_{aa}\left(1 - \left(\frac{F_{ab}}{F_{aa}}\right)\right) \qquad (12)$$

A positive value represents a red shift from A's location. A negative value represents a blue shift from A's location.

8)**Gravitational Time Contraction (Dilation) or Expansion:**

$$\Delta T_{aa} = T_{aa}\left(1 - \left(\frac{F_{ab}}{F_{aa}}\right)\right) \qquad (13)$$

A positive value represents gravitational time contraction (dilation) from A's location.
A negative value represents gravitational time expansion from A's location.

9)**The IRT procedure for determining the perihelion precession of Mercury without recourse to GRT is:**

a) Set up a coordinate system for the Sun and Mercury.

b) Use the IRT coordinate transformation equations to predict the future positions of the Sun and Mercury.

c) The perihelion shift of Mercury will be revealed when these future positions are plotted against time.

Also, the value of the shift can be determined from the plot.

Forces Based on Absolute Motions

The idea that absolute motion of interacting particles in the same direction gives rise to an attractive force, while absolute motion of interacting particles in the opposite directions gives rise to a repulsive force, is derived from the familiar electric current experiments in parallel wires. These experiments show that when electric currents are flowing in the wires in the same direction, the wires are attracted to each other, and when the currents are flowing in the opposite direction, the wires repel each other. Figures 1 and 2 illustrate these experiments graphically. The absolute motions of the electrons in the same direction cause a distortion in the E-Matrix that pulls the wires together--an attractive force. Conversely, the directions of absolute motion of the electrons in the opposite directions will cause a distortion in the E-Matrix that pulls the wires apart–a repulsive force.

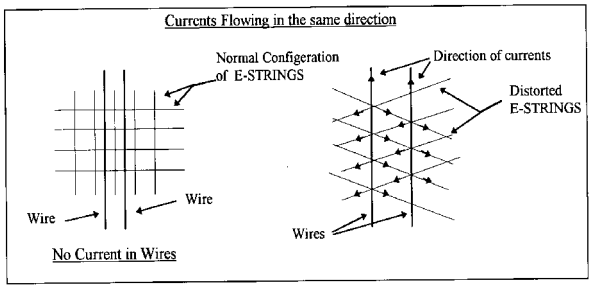


Figure 1. Currents (electrons) in the wires are flowing in the same direction, and therefore the force between the electrons is attractive. The right diagram that shows that the tension created in the E-Strings by the absolute motions of the electrons is pulling the wires together.

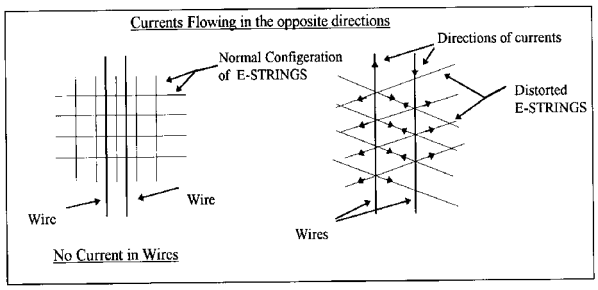


Figure 2. Currents (electrons) in the wires are flowing in the opposite direction, and therefore the force between the electrons is repulsive. The right diagram shows that the tension created in the E-Strings by the absolute motions of the electrons is pulling the wires apart.

Extending this interpretation of the electric-current experiments to include the orbiting motion of the S-Particles will enable us to explain all the nuclear forces between the interacting up quarks and down quarks [1,2]. This interpretation becomes the most important concept of Model Mechanics and it enables Model Mechanics to unite all the forces of nature naturally.

The CRE Force

Current physics posits that there are four forces of Nature: the electromagnetic force, the nuclear weak and strong forces, and gravity. Model Mechanics posits that there is a fifth force of Nature; the new force being the CRE force. As the name implies, the CRE force between any two objects is repulsive. While the CRE force is new to physical theory, it is not new to experience; it is what we commonly refer to as 'inertia'. In other words, the resistance between two objects to change their state of absolute motion is the CRE force between them. The CRE force between any two objects is always repulsive, and it is derived from the diverging structure of the E-Matrix.

To understand the CRE force, recall the inverse square law of physics. This law states that the intensity of light, gravity and electromagnetic force decreases with increasing distance r from the source is inversely proportional to r^2. The geometry of neighboring E-Strings emanating from any two objects also obeys the inverse square law. This means that each object will follow the diverging geometry of these neighboring E-Strings. Therefore, their path of motions in the E-Matrix will have a tendency to diverge from each other. This repulsive effect is identified as the CRE force. The CRE force between any two objects is not constant; it increases with the square of the distance between the objects. The CRE force is not the cosmological constant that Einstein inserted into his original GRT field equations. Although the cosmological constant is repulsive, it is not the CRE force predicted by Model Mechanics for the simple reason that it is constant.

The CRE force played an important role in the formation of our Universe, and is continuing to do so today. The repulsive CRE force, along with the

attractive electromagnetic force between gravitating objects shaped the primeval Universe into the Universe that we see today. The CRE force also played an important role in the manifestation of the nuclear weak force. Without the CRE force, there would be no nuclear weak force. It is the CRE force that initiates the radioactive decay of atoms. Perhaps, the most important function of the CRE force will be a role, in combination with the electromagnetic force, in the processes of life.

Model Mechanics predicted the repulsive CRE force in 1993. However, it was not discovered until 1998 when two independent groups of astronomers discovered that the Universe at the far reached regions is in a state of accelerated expansion. This observation is in direct conflict with the prediction of GRT. In order to explain this observation astronomers are now re-introducing the discarded repulsive Cosmological Constant to the GRT equation. The CRE force eliminates the need for this *ad hoc* approach.

Doppler Theory of Gravity (DTG)

Newton posited that gravity is a force, but he did not provide a mechanism for it. Newton's gravity model involved the unexplained phenomenon of action at a distance, which was troublesome for the physicists of his time. Also, Newton's equation for gravity was eventually found to be slightly inconsistent with observations. Recognizing the deficiencies in Newton's theory, Einstein formulated GRT, which is not a theory of force, but rather a theory of space-time, amounting to an extension of SRT to include gravity. IRT is a completed new theory of relativity. It includes SRT as a subset and its equations are valid in all environments…including gravity. It gives the same correct predictions for gravity as does GRT, but it avoids the following problematic predictions of GRT:

1) The expansion rate of the Universe as predicted by GRT does not match what is currently observed. GRT predicts that the expansion of the Universe is slowing down, and yet observation confirms that the expansion is speeding up.
2) The galactic rotational curves as predicted by GRT do not match those that are currently observed.
3) The path of travel of Pioneer 10 as predicted by GRT does not match what is observed.
4) GRT predicts the existence of black holes and singularities. If these absurd objects exist, they should be as abundant as the stars, and yet none them have been positively detected.
5) GRT fails to predict the existence of dark matter and dark energy.

Model Mechanics also gives rise to a new theory of gravity called Doppler Theory of Gravity (DTG). Like Newton's theory, DTG also treats gravity as a force but with an identified mechanism. Based on the provisions of Model Mechanics, the mechanism of gravity between two objects A and B moving in the stationary E-Matrix is as follows:

1) If both A and B are moving absolutely in the same direction, this gives rise to an attractive force because A's absolute motion distorts the surrounding stationary E-Matrix and B's absolute motion is confined to follow the distortion created by A; conversely, B's absolute motion distorts the surrounding stationary E-Matrix and A's absolute motion is confined to follow the distortion created by B.
2) The global structure of the stationary E-Matrix is divergent. Both A and B are confined to this global divergent structure as they travel in the stationary E-Matrix. This gives rise to the repulsive CRE force between A and B globally.

The force of gravity between A and B is the combined result of items (1) and (2). It is noteworthy that gravity is the sum of an attractive and a repulsive force acting on both A and B. This explains why the force of gravity is so weak compared to the electromagnetic and nuclear forces.

The above description for gravity suggests that the Newtonian equation for gravity can be modified to make it consistent with observations. The following is a modified Newtonian equation based on the above description for the force of gravity:

$$F = G * M_a M_b (j_a) \cdot (\pm j_b) / (r^2)(DF_a) \qquad (14)$$

F = The force of gravity between A and B as determined by A

G = Universal gravitational constant m^3/s^2*kg

M_a = Mass of object A in kg

M_b = Mass of object B in kg

$(j_a) \cdot (j_b)$ = Dot product of the directional vectors j_a and j_b. [Note: This dot product can be positive or negative.]

r = Distance in meters between A and B

DF_a = Doppler Factor as determined by A

$DF_a = F_{aa} / F_{ab}$

F_{aa} = Frequency of a standard light source in A's own frame as measured by A.

F_{ab} = Frequency of an identical standard light source in B's frame as measured by A. If F_{ab} is not constant, a mean value is used.

The dot product $(j_a)\cdot(j_b)$ in this new equation expresses the concept that not all objects in the Universe attract each other gravitationally. A positive dot product represents an attractive force, but a negative dot product represents a repulsive force. Those objects that have the same direction of absolute motion are attracted to each other, but those objects that have absolute motions in the opposite direction exert a repulsive force on each other. Assuming the Big Bang model is correct then the dot product of the vectors for all local regions of the Universe is +1. This means that gravity in the local region is attractive. The dot product for a distant region, say beyond the radius of the observable Universe, is -1. Therefore, gravity for all those distant regions is repulsive.

The Electromagnetic Force

This is the force observed between charged particles. It was determined that like-charged particles exert a repulsive force on each other while unlike charged particles exert an attractive force on each other. The reader will recall that a charged particle is the result of a clockwise or counterclockwise orbiting motion of its S-Particle around a specific E-String. A clockwise orbiting motion of the S-Particle gives rise to a positively charged particle. A counterclockwise orbiting motion of the S-Particle gives rise to a negatively charged particle. The charges between the interacting particles determine whether the force between them is attractive or repulsive. The following diagrams describe the electromagnetic force in Model Mechanical terms:

Interaction Between Negatively Charged Particles

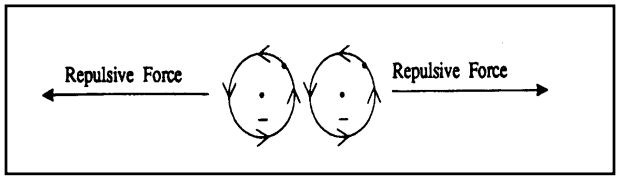


Figure 3. The force exerts on each other by two negatively charged particles. In this case, the S-Particles are traveling in the opposite directions and therefore the force between these particles is repulsive.

Interaction Between Positively Charged Particles

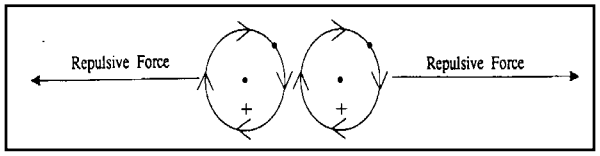


Figure 4. The force exerts on each other by two positively charged particles. In this case, the S-Particles are traveling in the opposite directions and therefore the force between the resulting particles is repulsive.

Interaction: Negatively and Positively Charged Particles

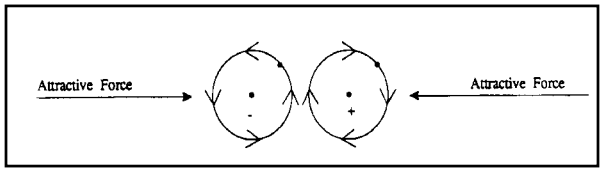


Figure 5. The force exerts on each other by a negatively and a positively charged particle. At the nearest point of approach the S-Particles are traveling in the same direction and therefore the force between them is attractive.

Note: The net attractive or repulsive force between any two interacting charged particles is not a constant force. The net force is determined by the direction of orbiting motions of their S-Particles at the closet point of approach. When the S-Particles are moving in the same direction at the closest point of approach then the net force between the charged particles is attractive. Conversely, when the S-Particles are moving in the opposite directions then the net force between the charged particles is repulsive. It is noteworthy to point out that the force between any two charged particles is alternating between attractive and repulsive for one complete orbit of their S-Particles. This property of the electromagnetic force is due to the fact that the direction of orbiting motions of the S-Particles is alternating between the same direction and opposite directions. This unique characteristic of the

electromagnetic force agrees with Maxwell's equation that the propagation of the electromagnetic force is alternating between the electric field and magnetic field.

The above diagrams illustrate how the electromagnetic force is manifested between charged particles. This force is long range because the distortions created in the E-Strings are long range. This description of the electromagnetic force eliminates the need for the complicated and abstractive quantum mechanical explanation. In addition, this explanation has no infinities to contend with because the electric charge is not within the particle itself. Therefore, there is no need for the dubious renormalization procedure to get rid of the infinities as in the quantum mechanical description of this force.

The Nuclear Strong Force

This force is responsible for binding the protons and the neutrons in the nucleus. At a more fundamental level, this force is responsible for the binding of the quarks of the protons and neutrons to form the nucleus. According to quantum mechanics the nuclear strong force is manifested by the exchange of messenger particles known as gluons.

The Model Mechanical description of the nuclear strong force is very simple. It is caused by the absolute motion (V_{suq} and V_{sdq}) of the S-Particles of the quarks in the protons and neutrons. This description of the nuclear strong force raises the question: Since the quarks in the protons and neutrons are negatively and positively charged particles, how do they manage to stick to each other? The answer is stacked-interaction. When two particles of the same charge are stacked on top of each other, their S-Particles are traveling in the same direction. Therefore, they exert an attractive force on each other. The following diagrams illustrate the stack interaction concept.

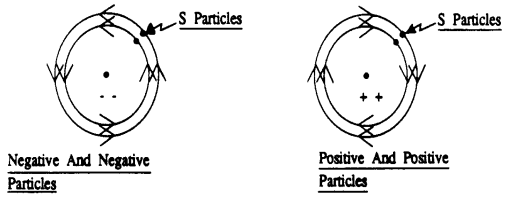


Figure 6. The stacked interactions of two similarly charged particles. The negative particles would be the down quarks and the positive particles would be the up quarks.

Note: All quarks of the same family have the same orbital diameter. The different orbital diameters shown here are served to illustrate the stacked-interactions. The negative and negative particle interaction is the stacked-interaction of the down quarks. The positive and positive interaction is the stacked-interaction of the up quarks.

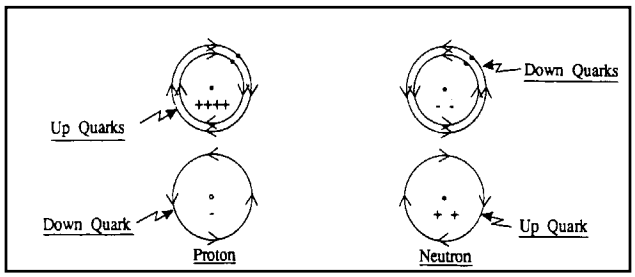


Figure 7. The stacked-interactions and the electromagnetic interactions in a proton and a neutron.

Note: The proton is formed by the stacked interaction of the up quarks and the electromagnetic interaction between the stacked up quarks and the down quark. The neutron is formed by the stacked interaction of the down quarks and the electromagnetic interaction between the stacked down quarks and the up quark.

It is noteworthy to point out that the attractive stacked-interactions are effective only within a short distance of 10^{-13} cm. At a greater distance than that the stacked-quarks exert a repulsive force on each other. This is the exact behavior of the nuclear strong force that we observed in the laboratory. Another peculiar property of the nuclear strong force is that it becomes stronger when the interacting particles are being pulled apart. This peculiar property is also predicted by Model Mechanics as follows: When the stacked particles are pulled apart the E-Strings surrounding them becomes more distorted. Therefore, the energy required to pull them further apart will be increased accordingly.

The Nuclear Weak Force

Quantum Mechanics describes this force as the force that causes the decaying processes of all the unstable particles through time. The quantum mechanical process for the weak force involves a process called the spontaneous breaking of symmetry. This process gives rise to the weak force messengers W^+, W^- and Z^0. These are virtual particles whose brief existence is financed by the uncertainty of energy and time relationship. Also, this description of the nuclear weak force depends on the existence of yet another class

of particles known as the Higgs particle. The Higgs particle is necessary because it is the mechanism that imparts mass to the weak force messengers.

Model Mechanics gives a much simpler description of the weak force. In the case of a heavy nucleus, such as a uranium nucleus, the decay is the result of the de-coupling of the stacked-interactions by a combination of neutron captures follow by the repulsive CRE force. The processes involved are as follows:

1. A free neutron is captured by a decaying nucleus
2. The stacked interactions at the site of neutron capture are weakened. This enables the repulsive CRE force to de-couple the weaken stacked-interactions and give rise to the nuclear weak force.
3. In the case of a subatomic particle, the decaying process is different. The best known subatomic particle-decaying process is the neutron decay, also known as the beta decay. Quantum Mechanics does not specify when a free neutron will decay or why it will decay in about sixteen minutes. On the other hand, Model Mechanics is capable of describing the neutron decay process in detail. The following diagrams will help the reader to visualize the processes involved.

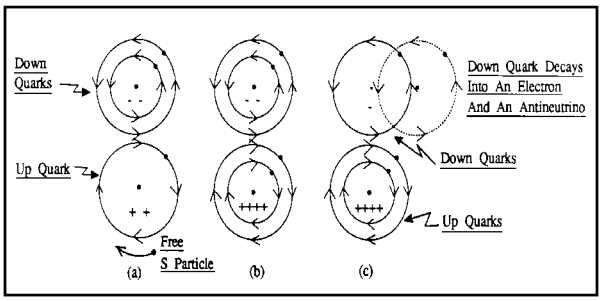


Figure 8. Schematic diagrams for the neutron decay process (Beta decay)

a) The up quark in an unbounded neutron exerts an attractive force on any free S-Particles that are traveling in the same direction as its S-Particle. When a free S-Particle follows the orbit of the orbiting S-Particle of the up quark, it becomes an up quark. This new up quark immediately forms a stacked interaction with the original up quark.

b) The down quark between the two-stacked up quarks is pulled closer to them because it feels the force from both of them.

c) This has the effect of moving the stacked down quarks laterally relative to each other. When the lateral movement is greater than the radius of the down quark, the force between the stacked down quarks becomes repulsive. This causes the down quark that feels less attractive force from the two stacked up quarks to peel away. The peel away down quark will then interact with a free S-Particle to give an electron and an antineutrino.

The decaying process for a subatomic particle such as a muon is different from that for a neutron. It was found that a muon at a speed closed to that of light would have a much longer decay length than that of a muon at the rest frame of the laboratory. When these decay lengths are converted to decay times they agree with the SRT time dilation equation. This led physicists to claim that the muon decaying process is a proof of the time dilation concept of SRT. The Model Mechanical explanation of the muon decay process is as follows:

1. The orbit of the muon's S-Particle is unstable and it will decay into a stable orbit of the electron. The muon takes the same amount of absolute time to decay in both cases. The different decay lengths observed are due to the different states of absolute motion of the muon.
2. When a muon is at the rest frame of the laboratory, its state of absolute motion is almost the same as that of the laboratory. Therefore, it is observed to have a very short decay length.
3. When the muon is accelerated to a speed close to that of light, the difference in the state of absolute motions between the accelerated muon and the laboratory muon is close to the speed of light. Therefore the accelerated muon is observed to have a very long decay length compared to the laboratory muon. This observed difference in decay lengths is interpreted as time dilation by current physics.

Model Mechanics Explains The Problematic Cosmological Observations

One of the most pressing problems of the Standard Big Bang Model is the observed horizon problem. The age of our universe is determined to be 14 billion years old in all directions and yet we observe the horizon for the opposite regions of our universe to be 28 billion years apart. In fact if all the regions are included the observed horizon of the universe is estimated to be 46 billion years. This means that these opposite regions of our universe cannot be in contact with each other at the Big Bang and this is known as the horizon problem.

Cosmologists invented the ad hoc *Inflation* hypothesis to explain the horizon problem. Model Mechanics explains the horizon problem naturally without resorting to the ad hoc *Inflation* hypothesis. The earth is in a state of absolute motion in the E-Matrix. This motion curves the E-Strings surrounding the earth. What we perceive as normal and straight E-Strings are actually severely curved E-Strings. In other words, when we look up in the sky we are actually receiving light from these curved E-Strings. This means that no matter what direction we look we are looking into the same curved E-Strings and thus the same region of the universe. This means that the perceived opposite regions of the universe are really the same region and therefore the perceived horizon problem was never existed. As it turns out, there is a perfect physical example of this phenomenon. The medical device gastro-scope made of fiber optics, allows a physician to examine the interior of a patient's stomach is such an example. No matter how the physician curves the eyepiece, he will still be seeing the same picture of the stomach.

In 1998 two independent groups of astronomers discovered that the far reached regions of the universe are in a state of accelerated expansion motion. This discovery is contrary to the predictions of GRT that predicts that the expansion of the universe should be slowing down. Astronomers revived the once discarded repulsive Cosmological Constant to explain the observed accelerated expansion. They posited that the universe is filled with a form of dark energy called Quintessence and this dark energy has the anti-gravity effect that gives rise to the Cosmological Constant. Model Mechanics predicted the accelerated expansion for those far reached regions of the universe in 1993. The basis for this Model Mechanical prediction is that gravity at those regions is repulsive with respect to us as described in the DTG equation. The repulsive CRE force of DTG can be considered as the dark energy posited by the astronomers.

Another problem arise from the GRT description of gravity is called the flatness problem. The flatness problem is that the observable universe appears to exist between an open and a closed universe. In an open universe, the matter density is less than the critical value and thus the gravitational braking effect is not able to halt the Big Bang expansion. This means that the universe will keep on expanding forever. In a closed universe the matter density is greater than the critical value and thus the gravitational braking effect will be able to halt the Big Bang expansion. This means that the universe will re-collapse before any galaxy would have time to form. In order for our universe to exist between an open and a closed universe the matter density must be fine tuned to be within one part in 10^{50} of the critical density value when the universe was a fraction of a second old. The inability of the Big Bang theory to explain why this degree of fine-tuning existed is what is known as the flatness problem. In Model Mechanics (DTG), gravity is the result of two gravitating objects having the same direction of absolute motions in the E-Matrix less the repulsive CRE force that exists between them. This description of gravity avoids the flatness problem completely.

The observed rotational curves of galaxies disagree with the predictions of GRT. These observed anomalous rotational curves correspond to curves for galaxies that are much more massive than the observed visible matters for these galaxies. The observed path of travel of the Pioneer 10 spacecraft disagrees with the predicted path given by GRT. Pioneer 10 was observed to be in a state of accelerated motion toward the sun. Astronomers explain both of these anomalous observations by claiming the existence of a dark matter in space although such an existence of dark matters is not within the framework of GRT or the Standard Model. Model Mechanics explains both of these anomalous observations by positing the existence of a dark matter in the form of free non-orbiting S-Particles.

Unification of Physics

Special Relative Theory (SRT) rejects the notion of absolute time and space. Also, SRT advanced the concept of energy and mass equivalency, which includes the idea that mass is convertible to energy and vice versa. These two concepts have been the foundation of theoretical physics developments for the past century. The results of these theoretical developments gave rise to the two pillars of modern physics: General Relativity Theory (GRT) and Quantum Mechanics (QM). GRT describes the large-scale universe including gravity while QM describes the microscopic universe including the electromagnetic and nuclear forces. Efforts to unify the electromagnetic force with the nuclear force have had some successes. The electro-weak theory unites the electromagnetic force with the nuclear weak force. However, this unification remains dubious because it depends on the

existence of a hypothetical particle called the Higgs particle. So far, physicists have not been able to find this hypothetical particle in the accelerators around the world. In fact, physicists at CERN are coming to the conclusion that the Higgs does not exist. There were also some successes in the effort of unifying the electro-weak force with the nuclear strong force. The resulting theory is called Grand Unification Theory (GUT). GUT predicts that the proton is not permanently stable. This led physicists around the world rushed to find evidence of proton decay. However, so far no evidence of proton decay was found. Attempts to unify gravity with the electromagnetic force and the nuclear forces were complete failures. There is no viable theory of quantum gravity.

The unification problems described above are the direct consequences of the foundations of modern physics, which deny the existence of absolute time and physical space (absolute space). The irony is that both GRT and QM contain math constructs that resemble physical space. The math construct space-time in GRT and the math construct field/virtual particle in QM are such examples. The difference between these math constructs and physical space is that the math constructs have no physical constraints. This lack of physical constraints leads to the infinity problems that plagued both GRT and QM. In GRT the lack of constraint leads to the infinity problems at the singularity where the theory breaks down completely. The other problem is that GRT gives no explanation why the force of gravity is capable of action-at-a-distance. In QM the lack of physical constraint leads to infinity problems during the formulation of the theories of electromagnetic and nuclear forces. This was especially true in the case of the theory of quantum electrodynamics (QED). In QED the electric change of a particle is resided within the particle. This leads to the infinity problems during the early development of QED. The infinity problem of QED was resolved by a dubious mathematical procedure called renormalization. A number of physicists, including Paul Dirac consider the renormalization technique a mathematical trick. He made the following comments during a lecture given in New Zealand in 1975:

> *"I must say that I am very dissatisfied with the situation, because this so-called 'good theory' does involve neglecting infinities which appear in its equations, neglecting them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves neglecting a quantity when it turns out to be small."*

Model Mechanics replaces the math constructs of space-time and field/virtual particle with the E-Matrix and the distortions or waves in the E-Matrix. It gives rise to the following postulates:

1) The E-Matrix is a stationary and structured light-conducting medium. It occupies all of pure space (pure void). It is comprised of very thin and elastic E-Strings and these E-Strings are repulsive to each other. There is an unknown compacting force that compresses these E-Strings together to form the E-Matrix.
2) The S-Particle is the only truly fundamental particle exists in our universe. The different orbiting motions of the S-Particles around the E-String(s) give rise to all the visible and stable particles in our universe.
3) All the processes of nature are the results of absolute motions of S-Particles or S-Particle systems in the E-Matrix.
4) All the forces of nature are the results of the S-Particle or S-Particle systems reacting to the distortions or waves in the E-Strings to which they are confined. The distortions or waves in the E-Strings, in turn, are the results of the absolute motions of the interacting S-Particles or S-Particle systems in the E-Matrix.
5) All the stable and visible matters are the results of orbiting motions of the S-Particles around specific E-Strings.

These postulates eliminate all the infinity problems that plagued both GRT and QM. It has the same mechanism for all the forces of nature and thus it unites all the forces of nature. It gives an explanation why the force of gravity is capable of acting at a distance. It explains the provisions of the Uncertainty Principle. It explains the weird results of all quantum experiments [3]. It eliminates the need for the undetectable force messengers in QM. It eliminates the need for the hypothetical and undetected Higgs particle. It explains the mass of a particle. It explains the charge of a particle. It leads to the discovery of the CRE force, which, in turn leads to a new theory of gravity. In short, Model Mechanics gives us a unique way to achieve the elusive goal of unifying all of physics.

Proposed Experiments To Detect Absolute Motions

Model Mechanics is based on the existence of the E-Matrix. Therefore absolute motions of objects in the E-Matrix should be detectable. However, numerous past attempts to detect absolute motion were failures. The most notable of these is the Michelson-Morley Experiment (MMX) [3]. In this experiment a light beam was split into two parts that were directed along the two arms of the instrument at right angles to each other, the two beams being reflected back to recombine and form interference fringes. Any shift in the interference fringes as the apparatus is rotated would mean the detection of absolute motion of the apparatus. To everyone's chagrin, the MMX produced a null result. However, the MMX null result does not mean that there is no absolute motion of the apparatus. In their interpretation of the MMX null result Michelson-Morley failed to ask the relevant question: What is the direction of absolute motion of the apparatus with respect to the defined horizontal plane of the light rays that will produce a null result for all the orientations of the horizontal arms? The answer to this question is: If the apparatus is moving vertically then a null result will be obtained for all the orientations of the horizontal arms. What this mean is that the MMX as designed is not capable of detecting the absolute motion of the apparatus. In order to detect absolute motion using the MMX, the plane of the arms must be oriented vertically. This conclusion is supported by the observed gravitational red shift (gravitational potential) in the vertical direction.

The new interpretation of the MMX null result gives rise to a new concept for the propagation of light as follows:

How does light get from point A to point B? The current assumption is that, locally, light travels in a straight line towards the target, and that, in a train of light pulses, the first pulse hits the target is the first one the source generated. These assumptions both make sense if the target is stationary relative to the light pulses, but if the target moves the second assumption could be erroneous. Figure 9 describes a thought experiment that is currently used by physicists to derive the time dilation equation. A light clock is constructed of two mirrors parallel to each other with light pulses bouncing between them. In one period of the clock, a light pulse travels up to the top mirror and returns back to the bottom mirror. The diagram shows that the light pulse is presumed to travel a slant path when the light clock is in motion. This is not a realistic description of the actual event. It raises the question: How does light know when to follow a vertical path and when to follow one of the infinite numbers of slant paths? It is more realistic to say that light will always follow the perpendicular path on its way to the upper mirror. The reason is that the vertical path is the direction where all the light pulses are directed. Figure 10 shows this: the first pulse of a train of pulses follows the original path AB, but the pulse detected at "E" travels the path CE.

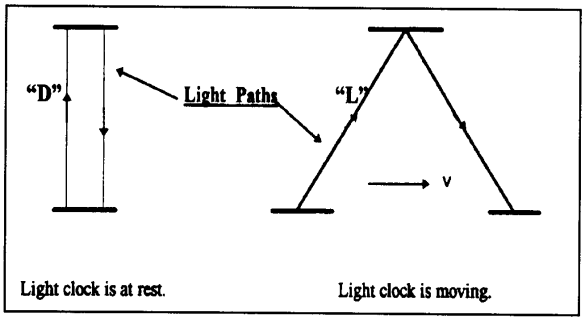


Figure 9. Light paths in a light clock at rest and in motion.

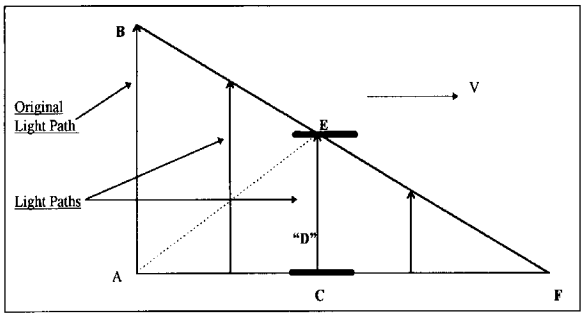


Figure 10. Current physics says that AE is the path that light follows to the upper mirror and theangle of this path is depended on the length AC that is depended on the speed of the light clock.

With this description of the light paths, the first pulse is never detected at "E." The light pulse detected at "E" is generated by the source at a later time. It turns out that this description of light paths is also capable of giving us the time dilation equation by using the Pythagorean theorem. The reason is that the original light path (AB) is equal to the assumed light path (AE) and both are the radii of a light sphere at the point of origin "A". It is noteworthy that as the speed of the mirrors approaches light speed a light pulse will take a longer time to reach the upper mirror. When the mirrors are moving at the speed of light, no light pulse is able to reach the upper mirror at all. Current physics interprets this situation as time standing still at the speed of light. The new interpretation is that time keeps on ticking at

all speeds of the light clock. The amount of time (duration) passed depends on the length of the original light path AB divided by the speed of light 'c'. This new interpretation suggests that absolute time for a moving frame is not slowed or dilated as currently assumed. The specific amount of absolute time (duration) required for light to travel the original light path AB is equal in all frames. A light clock runs slow when it is in motion because it is not catching the first light pulses, but rather some later one. The lower elapsed time recorded by a moving clock because the passage of time is not fully detected when the clock is in a state of motion.

The new interpretation of the MMX null result and the new concept for the propagation of light enable us to design the following experiments to detect absolute motion:

Experimental Set Up:

1) Two sets of cesium clocks A1, A2 and B1, B2 are located at the middle of a 120 meters long straight rail track. Distances of 25 meters and 50 meters on both sides of the mid-point are marked off with a physical ruler.
2) Each set of clocks is equipped with a laser light sources and a beam splitter that splits the laser beam into two continuous beams. One beam goes to detector "A" and the other goes to detector "B".
3) Each set of clocks is equipped with a shutter that allows the two laser beams to pass through it for any desired time intervals.
4) Each set of clocks is equipped with a circular surface detector and the detecting surface can vary from 3 mm to 20 cm in diameter.
5) Each set of clock is equipped with a reflecting mirror.
6) A1 and B1 are not running. A2 and B2 are synchronized and running.

Experiment Group #1: To Detect The Absolute Motion Of The Distant Clock At 50 Meters

a) Move both sets of clocks simultaneously in the opposite directions at a rate of 10 meters/day (1 day = 86,400 seconds) and stop them at the 25 meters marks (after 2.5 days). The clocks are now 50 meters apart.
b) Both detecting surfaces are set at 3mm in diameter.
c) Do the following experiments from A's location.
d) A trial of the experiment is consisted of an opening and closing of the shutter for a specific time interval. The following trials at the following time intervals are made: 1 second, 2 seconds, 3 seconds, 4 seconds, 5 seconds, 6 seconds, 7 seconds, 8 seconds, 9 seconds and 10 seconds. The trials are conducted from A's location.
e) Laser beam A will activate and de-activate clock A1 for each trial and the results are identified as T'a1, T'a2, T'a3, T'a4, T'a5, T'a6, T'a7, T'a8, T'a9 and T'a10.
f) Laser beam B will activate and de-activate clock B1 for each trial and the results are identified as T'b1, T'b2, T'b3, T'b4, T'b5, T'b6, T'b7, T'b8, T'b9 and T'b10.
g) The difference in activation time between clocks A1 and B1 for each trial is identified as follows: ΔT'1, ΔT'2, ΔT'3, ΔT'4, ΔT'5, ΔT'6, ΔT'7, ΔT'8, ΔT'9, and ΔT'10.
h) Increase the detecting surface to 20 cm in diameter then perform a trial using the 1-second time interval to establish that there is no difference in activation time between A1 and B1 for this large detecting surface. Now reduce the diameter of the detecting surface gradually to find the diameter where the activation time between A1 and B1 start to show a difference. Call this critical diameter D_{50}.
i) Cover the detecting surface completely with a 20 cm diameter dish. A slit of 2mm wide is cut from the center of the dish to the outer rim of the dish. Slowly rotate the dish to find the direction of absolute motion of the detector. That direction is evident when the slit is in line with the direction of absolute motion of the detector and activates the clock B1 for the same amount of time as the shutter opening and closing at A's location.
j) Repeat the above experiments from the "B" location.

The SRT Predictions For Group #1 Experiments:

- The activation time for the B1 clock is the same as that for the A1 clock for all trials.
- The difference in activation time between A1 and B1 is zero for each trial. ΔT'1=ΔT'2=ΔT'3=ΔT'4=ΔT'5=ΔT'6=ΔT'7=ΔT'8 =ΔT'9=ΔT'10=0.

- Increase the diameter of the detecting surface will have no effect on activation time on the B1 clock for each trial.
- There is no absolute motion of clock B1 and therefore there is no direction of absolute motion.
- Repeating the above experiments from the B location will get the same results as above.

The Model Mechanical Predictions For Group #1 Experiments:

- The activation time for the B1 clock is less than that for the A1 clock for each trial. This is due to the B1 clock is in a state of absolute motion in the vertical direction while the laser is in transit from A to B.
- The difference in activation time between A1 and B1 is the same for each trial and it is greater than zero.
- Increase the diameter of the detecting surface will bring the activation time for the B1 clock equal to that of the A1 clock.
- The absolute motion of the clock B1 (V_{50}) can be calculated using the following equation:

$$V_{50} = \frac{D_{50}}{2\Delta T1} \qquad (15)$$

- The direction of absolute motion of the B1 clock is vertical.
- Repeating the above experiments from the "B" location will get the same results as above.

Experiment Group #2: To Measure the One-Way and Two-Way Speed of Light at 50 Meters Apart

a) The clocks A2 and B2 are 50 meters apart and are still synchronized according to SRT and Model Mechanics.
b) Measure the one-way speed of light using clocks A2 and B2 from the "A" location.
c) Measure the one-way speed of light using clocks B2 and A2 from the "B" location.
d) Measure the two-way speed of light using clock A2.
e) Measure the two-way speed of light using clock B2.

The SRT Predictions For Group #2 Experiments:

- The one-way speed of light is c as measured from the "A" location.
- The one-way speed of light is c as measured from the "B" location.
- The one-way speed of light is isotropic.
- The two-way speed of light is c using clock A2.
- The two-way speed of light is c using clock B2.
- The two-way speed of light is isotropic.

The Model Mechanical Predictions For Group #2 Experiments:

- The value for the one-way speed of light is less than c as measured from the "A" location.
- The value for the one-way speed of light is less than c as measured from the "B" location.
- The one-way speed of light is isotropic. In other words, the value for the one-way speed of light from A→B is equal to from B→A.
- The calculated value for the one-way speed of light can be make to equal to c by reducing the measured flight time by a factor of (ΔT'1).
- The two-way speed of light is c using clock A2.
- The two-way speed of light is c using clock B2.
- The two-way speed of light is isotropic.

Experiment Group #3:To Detect The Absolute Motion Of The Distant Clock At 100 Meters

a) Move both sets of clocks at the 25 meters marks simultaneously in the opposite directions at a rate of 10 meters/day and stop them at the 50 meters marks (after 2.5 days). The clocks are now 100 meters apart.
b) Both detecting surfaces are set at 3mm in diameter.
c) Do the following experiments from A's location.
d) A trial of the experiment is consisted of an opening and closing of the shutter for a specific time interval. The following trials at the following time intervals are made: 1 second, 2 seconds, 3 seconds, 4 seconds, 5 seconds, 6 seconds, 7 seconds, 8 seconds, 9 seconds and 10 seconds. The trials are conducted from A's location.
e) Laser beam A will activate and de-activate clock A1 for each trial and the results are identified as T"a1, T"a2, T"a3, T"a4, T"a5, T"a6, T"a7, T"a8, T"a9 and T"a10.
f) Laser beam B will activate and de-activate clock B1 for each trial and the results are identified as T"b1, T"b2, T"b3, T"b4, T"b5, T"b6, T"b7, T"b8, T"b9 and T"b10.
g) The difference in activation time between clocks A1 and B1 for each trial is identified as ΔT"1,

ΔT"2, ΔT"3, ΔT"4, ΔT"5, ΔT"6, ΔT"7, ΔT"8, ΔT"9, and ΔT"10.

h) Increase the detecting surface to 20 cm in diameter then perform a trial using the 1-second time interval to establish that there is no difference in activation time between A1 and B1 for this large detecting surface. Now reduce the diameter of the detecting surface gradually to find the diameter where the activation time between A1 and B1 start to show a difference. Call this critical diameter D_{100}.

i) Cover the detecting surface completely with a 20 cm diameter dish. A slit of 3mm wide is cut from the center of the dish to the outer rim of the dish. Slowly rotate the dish to find the direction of absolute motion of the detector. That direction is evident when the slit is in line with the direction of absolute motion of the detector and activates the clock B1 for the same amount of time as the shutter opening and closing at A's location.

j) Repeat the above experiments from the "B" location.

The SRT Predictions For Group #3 Experiments:

- The difference in activation time between A1 and B1 is zero for each trial. ΔT"1=ΔT"2=ΔT"3=ΔT"4=ΔT"5=ΔT"6=ΔT"7=ΔT"8=ΔT"9=ΔT"10=0.
- Increase the diameter of the detecting surface will have no effect on activation time on the B1 clock for each trial.
- There is no absolute motion of clock B1 and therefore there is no direction of absolute motion.
- Repeating the above experiments from the B location will get the same results as above.

The Model Mechanical Predictions For Group #3 Experiments:

- The activation time for the A1 clock is greater than that for the B1 clock for each trial. This is due to the B1 clock is in a state of absolute motion in the vertical direction while the laser is in transit from A to B.
- The difference in activation time between A1 and B1 is the same for each trial.
- ΔT"1=ΔT"2=ΔT"3=ΔT"4=ΔT"5=ΔT"6=ΔT"7=ΔT"8=ΔT"9=ΔT"10
- Increase the diameter of the detecting surface will bring the activation time for the B1 clock equal to that of the A1 clock.
- The absolute motion of the clock B1 (V_{100}) can be calculated using the following equation:

$$V_{100} = \frac{D_{100}}{2\Delta T'1} \qquad (16)$$

- The direction of absolute motion of the B1 clock is vertical.
- Repeating the above experiments from the "B" location will get the same results as above.

Experiment Group #4: Measure the One-Way and Two-Way Speed of Light at 100 meters

a) The clocks A2 and B2 are 100 meters apart and are still synchronized according to SRT and Model Mechanics.

b) Measure the one-way speed of light using clocks A2 and B2 from the "A" location.

c) Measure the one-way speed of light using clocks B2 and A2 from the "B" location.

d) Measure the two-way speed of light using clock A2.

e) Measure the two-way speed of light using clock B2.

The SRT Predictions For Group #4 Experiments:

- The one-way speed of light is c as measured from the "A" location.
- The one-way speed of light is c as measured from the "B" location.
- The one-way speed of light is isotropic.
- The two-way speed of light is c using clock A2.
- The two-way speed of light is c using clock B2.
- The two-way speed of light is isotropic.

The Model Mechanical Predictions For Group #4 Experiments:

- The value for the one-way speed of light is less than c as measured from the "A" location.
- The value for the one-way speed of light is less than c as measured from the "B" location.
- The one-way speed of light is isotropic. In other words, the value for the one-way speed of light from A→B is equal to from B→A.
- The calculated value for the one-way speed of light can be make to equal to c by reducing the measured flight time by a factor of (ΔT"1).

- The two-way speed of light is c using clock A2.
- The two-way speed of light is c using clock B2.
- The two-way speed of light is isotropic.

Conclusions:

Model Mechanics enables us to describe all the processes and interactions in terms of absolute motion of S-Particle or S-Particle systems (a new description of matter) in the E-Matrix (a new description of physical space). The unique structure of the E-Matrix led to the discovery of the CRE force, a repulsive component of gravity. The CRE initiates the nuclear weak force and participates in all the processes of nature. The absolute motion of S-Particles or S-Particle systems in the E-Matrix in the same direction gives rise to an attractive force, while absolute motion of S-Particle or S-Particle systems in the opposite directions gives rise to a repulsive force. Gravity is the attractive force between two objects having the same direction of absolute motion in combination with the repulsive CRE force creates by the same absolute motion of the objects. This explains why the force of gravity is so weak compared to the other forces. The electromagnetic force and the nuclear weak and strong forces are also the result of absolute motion between the interacting S-Particles. These forces are the results of the interacting particles reacting to the distortions or waves in the E-Strings to which they are confined. A candidate unification of all physics is achieved with this simple description of the current universe.

Model Mechanics gives new interpretations for the MMX null results and that, in turn, leads to a new concept for the propagation of light. This enables us to design viable experiments to detect absolute motion. Performing these designed experiments will confirm the existence of the E-Matrix frame and will also provide a way to unify all of physics.

References

[1] Seto KH. Unification of Physics. Journal of the Theoretics. pp 5-11. **http://www.journaloftheoretics.com/Links/Papers/Seto.pdf**

[2] Seto KH. The Physics of Absolute Motion. ISBN0-9647136-1-6, KHS Publishing, pp 57-70.
http://www.erinet.com/kenseto/book.html

[3] Michelson AA, Morley EW. On the Relative Motion of the Earth And the Luminiferous Ether. Am J Sci. Third Series 1887;34:333-45.

Crohn's Disease

Gail Michael

Chicago, the United States, Beacon1415@aol.com

Abstract: Crohn's disease is caused by someone the crohnie knows well and is in a relationship with (friend or family) who happens to be on a stimulant drug, such as amphetamine, diet drug, anti-depressant, marijuana, cocaine, depakaote, buspar, seratonin drug. The med or drug user is being accused of being a drug junkie. [Nature and Science. 2005;3(2):21].

Key words: Crohn's disease

Crohn's disease is caused by someone the crohnie knows well and is in a relationship with (friend or family) who happens to be on a stimulant drug, such as amphetamine, diet drug, anti-depressant, marijuana, cocaine, depakaote, buspar, seratonin drug or 1 of many I could not know of that have an anti-depressant in it as an ingredient. When that person is on the drug, he/she can unknowingly cause a vulnerable person to develop the symptoms of Crohn's. The strange and unbelievable part that defies logic is that the two (or more) persons do not have to be in the same room for the crohnie to have the symptoms. They can be miles and miles apart. All it takes is the mind connection and the drug. Asking the person to cease the drug and substitute it with a tranquilizer such as Valium or Ativan rarely will be successful, as the person cannot believe that what they use or ingest can harm another person.

In that case, I cease ALL contact with the person, hoping that time will diminish and negate our thoughts of each other, because as long as the drug and thoughts continue, so do the symptoms. Much like a divorce, it takes time to release the thoughts.

As for children, younger than 16 or so, they rarely have friends that are on a drug (or med, if that sounds better) so usually it is a caretaker or parent that may be on an anti-depressant or one of the other meds, not thinking it important to tell the child they are on the med. May I add, this is not to say discontinue treatment with the physician, until the child or person is completely feeling normal.

The latest med I discovered that has an antidepressant in it is Flexeril, an anti-spasmodic, by looking in the PDR (Physicians Desk Reference) and who would ever guess.

To cite something I was told by a pediatrician at the County Hospital in Chicago of the United States, is that they have babies of drug addicted mothers that throw up constantly. The doctor said they throw up because they are trying to bond with the mother and cannot because she is on a hard drug. So they remove the baby from the mothers care and have a surrogate mother take over.

The vomiting illness is called rumination. Their action is correct, but their reasoning does not coincide with mine. I tried to explain, the baby IS bonded with the mother but is reacting by throwing up BECAUSE of the stimulant drug the mother is on.

Whether the idea was accepted or not, I cannot say. I realize it is a difficult concept to understand. However, if one can identify with this theory, or even if not, ask around and not be bashful, the results will be the cessation of Crohn's and UC, if not immediately, then in the time that is needed to break the bond. Some misconception about this theory.

The thoughts have to be bad thoughts. No, usually they are good thoughts or any thoughts, because one tends to forget the person if they do not like them.

The med or drug user is being accused of being a drug junkie. Some are illegal drugs, but most are the commonly used drugs prescribed by physicians. The person that has Crohn's only a short time has an advantage, because no drugs or surgeries complicate the matter. In a long standing illness, there may be fallouts from the past treatments, but the Crohn's will cease and no further new symptoms will occur. This is not moral issue, but a health issue to try get someone well.

References

[1] http://digestive.niddk.nih.gov/ddiseases/pubs/crohns.
[2] http://www.angelfire.com/ga/crohns.

Application of ANN Model Based Simulated Annealing to Analyzing the Rice Flood Loss

Qiang Fu[1, 2], Nan Sun[2], Wei Zu[2]

1. Doctoral Working Station of Beidahuang Company, General Bureau of Agricultural in Heilongjiang, Harbin, Heilongjiang 150040, China; fuqiang100@371.net
2. School of Water Conservancy & Civil Engineering, Northeast Agriculture University, Harbin, Heilongjiang 150030, China

Abstract: Through improving the error back propagation artificial neural networks, a kind of new model of back-propagation artificial neural networks based on simulated annealing algorithm was built. The new model can overcome the disadvantage of traditional BP, which are slower convergence velocity, lower accuracy and easily fall into the local optimum. Through using the new model to analyzing the rice flood loss, it was showed that there are several advantages such as fast convergence velocity, good whole optimum, and strong practicability. The new improved model can be used in some other correlative research area. [Nature and Science. 2005;3(2):22-27].

Key Words: simulated annealing (SA); artificial neural networks (ANN); flood loss

1 Preface

In the flood disaster evaluation system, the model of flood disaster loss r is a response function to the factors and environment of disaster. The model has an important position in flood evaluation. It can be used in many aspects during the course of monitor natural disaster due to the nonlinearity, dynamic and multidimensional. So, building the suitable model become difficulty and hotspot in the research area of flood disaster.

In order to study the model and its application, the author take the rice disaster loss caused by flood as an example. Thus, the back-propagation artificial neural networks based on simulated annealing algorithm has been used.

It shows that the yield of rice is influenced by different growth stage, inundated deep, inundated diachronic (Zhu, 1990). The model of rice flood loss is to study the relationship between the rate of rice reduction of output and some influenced factors. The research work has great significance to ascertain the standard of water logging control, seek optimize control model, and improve economic benefit of reducing disaster reasonably. Obviously, neural networks model is an efficient approach to settle the approximating function. Jin put forward a kind of networks and it is feasible to built the rice flood loss model，but the training time of the networks is excessively long (Jin, 2000). Thus, a new method of training the networks named SA was put forward to improve the original networks.

2 Back Propagation Networks Based on Simulated Annealing Algorithm Model (SABP)

Because simulated annealing algorithm originates from annealing process and simulation it adopts Metropolis rule, which can avoid the local optimizing trap (Li, 2000). It can be combined with BP model. Thus, a new algorithm of BP named SABP (back propagation networks based on simulated annealing algorithm) has been put forward. When BP algorithm run into local optimization, SA algorithm can help BP model to handle the problem. Therefore, an arithmometer should be designed.

2.1 SABP Modeling

First, a function of SABP should be defined. That is FUNCTION SABP (W, B). The steps of modeling are as follows.

Step 1: Initialization weight matrix ($W^{(0)}$) and threshold value matrix ($B^{(0)}$).

Step 2: Initialization parameter ($\xi > 0$) of precision controls and preset circle times ($M > 0$).

Step 3: Define the initial error $SSE(k) = \xi + 0.5$.

Step 4: While $SSE(k) > \xi$ and $n < M$, then,

(1) Initialization the total error. Let $E = 0$, $n = 1$, $Count = Const$, $Number = 0$.

(2) Select a pair of sample random (X_p, Y_p) . Firstly, the practical networks output O_p corresponding to X_p should be calculated. Secondly, the error of output layer E_p should be calculated. Thirdly, calculate the accumulative total error $E = E + E_p$. Fourthly, the new weight value $W^{(1)}$ and new threshold value $B^{(1)}$ can be calculated by rectifying. At last, let $p = P - 1$ and $n = n + 1$. Then, if p equals zero, the program is over. Otherwise, the program should calculate the next pairs of sample.

(3) Calculate the output error of all samples. $SSE(k+1) = E/2$. If $SSE(k+1) < \xi$, the precision is satisfied. Otherwise, the program will run into the next step to train the networks continually. It cannot stop until the precision is satisfied.

(4) If border upon two circle total error $\Delta SSE = abs(SSE(k+1) - SSE(k)) < \varepsilon$, then set $Number = Number + 1$; When $Number = Const$, the Simulated Annealing Algorithm should be used. $SSE(k+1) = SA(W, B)$.

2.2 SA Algorithm of Simulated Annealing Algorithm

FUNCTION SSE = $SA(W, B, SSE, P, O)$. Where, W, B, SSE, P, O are weight value, threshold value matrix, input sample and expectation output sample.

Step 1: set original temperature ($T(0) = T_0$) and, iteration time ($k = 0$).

Step 2: According to the certain random process, a little-sized weight value disturbance σ will be gained. The weight value and threshold value will be revised.

$W1 = W + \sigma$; $B1 = B + \sigma$. Then, the total error of the networks $E1$ can be calculated.

Step 3: If $E1 < SSE$, let $SSE = E1, W = W1, B = B1$.

Step 4: If $E1 \geq SSE$, it should be revised according to Metropolis rule.

At first，calculate probability $p = \exp[-\Delta E / kT(t)]$, $T(t)$ is temporal temporality, then bring a random decimal fraction $\gamma = random(0,1)$, if $p > \gamma$, receive this revise： $SSE = E1$, $W = W1$, $B = B1$; otherwise, estimate weight value disturbance time, which whether exceed preset value or not，if not transfer step 2.

Step 5: $t = t + 1$, let $T = T_0 / \log(1 + t)$.

Step 6: If the temperature falls into the minimum temperature, then the total error cab is calculated. $SSE = E1$; IF the temperature hasn't reached the minimum temperature, the program runs into step 2.

In order to handle the disadvantage of original method of BP-ANN, such as slower convergence speed and longtime of training, the author put forward a new method to handle the problem (Fu, 2002; Zhou, 2000).

The improved BP algorithm named SABP.

3 Application of the SABP Algorithm

Rice yield is influenced by different generating process, inundated deep and inundate last a period time. The model of rice flood loss is a mapping problem of multi-dimension and nonlinear function. So, it is suitable to handle by BP-ANN model.

Table 1 shows the result of reduction in yield of rice (Zhou, 2000). In Table 1, when the first column value is one, it represents rice early tillering stage. When the second column value is one, it represents rice middle of tillering. When the third column value is one, it represents rice late tillering stage. When the forth column value is one, it represents rice booting stage. When the fifth column value is one, it represents rice spiking and blooming stage. When the sixth column value is one, it represents rice milk stage. When the seventh column value is one, it represents stage of yellow ripeness. The eighth column value represents rice flooded period (unit：day), the ninth column value represents quotient of flooded deep to plant height, and the tenth column value represents lapserate. The positive one shows rice increasing in crop yield and the negative one shows rice drop in crop yield. There are 105 groups of training samples, the last seven rows represent test sample.

Thus, the nerve cell of ANN input layer is 9，that are from the first column to the ninth column in Table 1. The output nerve cell is 1 and the nerve cell in hidden layer 1s 14. The network has three layers. At last, the

frame of networks is 9-14-1.

Through training 3213 times, the networks convergence. Now, the minimum of network energy function is 7.2953×10^{-4}. The total error of all the 105 groups of training sample is 309.201 and the average error is 2.945. The output value of each sample sees also to the eleventh column in the Table 3—1. The last seven rows represent test sample, and the network forecast overall error is 19.538. The data of the twelfth column is the results of reference calculated by MAGA-BP Algorithm (Zhou, 2000). The total error summation is 133.823, average error is 1.2475 and the total error of forecasting sample t is 21.037 (Table 1, Table 2).

Table 1. Result of reduction in yield of rice analyzed by SABP

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	SABP	MAGABP
1	0	0	0	0	0	0	2	0.50	-5.850	-7.229	-5.842
1	0	0	0	0	0	0	2	0.75	-8.440	-8.534	-8.923
1	0	0	0	0	0	0	2	1.10	-13.490	-12.430	-16.115
1	0	0	0	0	0	0	4	0.50	-9.730	-4.704	-8.047
1	0	0	0	0	0	0	4	0.75	-10.130	-15.952	-10.255
1	0	0	0	0	0	0	4	1.10	-27.170	-28.275	-28.925
1	0	0	0	0	0	0	6	0.50	-9.100	-11.987	-9.316
1	0	0	0	0	0	0	6	0.75	-39.570	-31.829	-32.719
1	0	0	0	0	0	0	6	1.10	-55.670	-52.800	-55.635
1	0	0	0	0	0	0	8	0.50	-17.210	-17.688	-17.533
1	0	0	0	0	0	0	8	0.75	-33.260	-39.846	-29.488
1	0	0	0	0	0	0	8	1.10	-72.220	-74.058	-70.689
1	0	0	0	0	0	0	10	0.50	-18.460	-17.254	-13.239
1	0	0	0	0	0	0	10	0.75	-42.690	-38.738	-37.287
1	0	0	0	0	0	0	10	1.10	-89.950	-89.775	-90.547
0	1	0	0	0	0	0	2	0.50	-5.680	-7.167	-5.708
0	1	0	0	0	0	0	2	0.75	-15.710	-12.510	-11.559
0	1	0	0	0	0	0	2	1.10	-20.100	-21.654	-20.278
0	1	0	0	0	0	0	4	0.50	-10.290	-8.419	-7.335
0	1	0	0	0	0	0	4	0.75	-11.870	-13.097	-9.696
0	1	0	0	0	0	0	4	1.10	-15.450	-16.034	-15.465
0	1	0	0	0	0	0	6	0.50	-3.140	-8.637	-8.664
0	1	0	0	0	0	0	6	0.75	-9.710	-7.179	-9.874
0	1	0	0	0	0	0	6	1.10	-18.740	-14.793	19.497
0	1	0	0	0	0	0	8	0.50	-3.740	-2.013	-5.337
0	1	0	0	0	0	0	8	0.75	-1.530	-4.653	-1.986
0	1	0	0	0	0	0	8	1.10	-36.120	-39.584	-36.178
0	1	0	0	0	0	0	10	0.50	-3.320	-1.262	-9.932
0	1	0	0	0	0	0	10	0.75	-23.470	-26.035	-14.237
0	1	0	0	0	0	0	10	1.10	-98.440	-96.329	-95.241
0	0	1	0	0	0	0	2	0.50	0.060	-0.972	0.061
0	0	1	0	0	0	0	2	0.75	-0.020	7.508	-0.470
0	0	1	0	0	0	0	2	1.10	-5.080	-7.791	-4.198
0	0	1	0	0	0	0	4	0.50	19.040	12.171	-20.118
0	0	1	0	0	0	0	4	0.75	14.520	10.485	-11.759

0	0	1	0	0	0	0	4	1.10	-31.410	-28.066	-25.938
0	0	1	0	0	0	0	6	0.50	-3.240	1.887	-3.716
0	0	1	0	0	0	0	6	0.75	-5.510	-5.042	-5.512
0	0	1	0	0	0	0	6	1.10	-53.960	-56.206	-55.211
0	0	1	0	0	0	0	8	0.50	-16.700	-10.158	-12.486
0	0	1	0	0	0	0	8	0.75	-15.720	-22.174	-15.375
0	0	1	0	0	0	0	8	1.10	-84.080	-81.796	-82.198
0	0	1	0	0	0	0	10	0.50	-6.860	-11.627	-7.222
0	0	1	0	0	0	0	10	0.75	-36.870	-34.768	-37.745
0	0	1	0	0	0	0	10	1.10	-96.500	-98.573	-91.044
0	0	0	1	0	0	0	2	0.50	-47.760	-46.079	-47.637
0	0	0	1	0	0	0	2	0.75	-69.670	-71.001	-66.843
0	0	0	1	0	0	0	2	1.10	-77.150	-76.756	-74.652
0	0	0	1	0	0	0	4	0.50	-39.390	-42.77	-39.393
0	0	0	1	0	0	0	4	0.75	-76.070	-71.885	-75.284
0	0	0	1	0	0	0	4	1.10	-86.550	-85.137	-84.257
0	0	0	1	0	0	0	6	0.50	-69.400	-65.816	-68.732
0	0	0	1	0	0	0	6	0.75	-79.610	-88.125	-81.892
0	0	0	1	0	0	0	6	1.10	-95.740	-95.376	-94.547
0	0	0	1	0	0	0	8	0.50	-85.900	-84.916	-86.575
0	0	0	1	0	0	0	8	0.75	-97.100	-97.255	-93.809
0	0	0	1	0	0	0	8	1.10	-97.750	-99.645	-94.643
0	0	0	1	0	0	0	10	0.50	-86.370	-89.185	-82.263
0	0	0	1	0	0	0	10	0.75	-96.330	-95.038	-94.403
0	0	0	1	0	0	0	10	1.10	-100.00	-96.533	-94.675
0	0	0	0	1	0	0	2	0.50	-11.600	-10.350	-11.881
0	0	0	0	1	0	0	2	0.75	-40.780	-47.301	-40.601
0	0	0	0	1	0	0	2	1.10	-75.960	-78.550	-72.944
0	0	0	0	1	0	0	4	0.50	-10.260	-3.701	-10.726
0	0	0	0	1	0	0	4	0.75	-79.340	-65.618	-78.117
0	0	0	0	1	0	0	4	1.10	-98.200	-97.073	-93.855
0	0	0	0	1	0	0	6	0.50	-3.020	-25.760	-1.584
0	0	0	0	1	0	0	6	0.75	-85.710	-84.509	-86.281
0	0	0	0	1	0	0	6	1.10	-100.00	-103.860	-94.347
0	0	0	0	1	0	0	8	0.50	-71.430	-51.687	-71.022
0	0	0	0	1	0	0	8	0.75	-90.360	-92.589	-92.177
0	0	0	0	1	0	0	8	1.10	-100.00	-100.370	-94.567
0	0	0	0	1	0	0	10	0.50	-67.660	-75.256	-77.426
0	0	0	0	1	0	0	10	0.75	-100.00	-99.615	-93.895
0	0	0	0	1	0	0	10	1.10	-100.00	-98.539	-94.649
0	0	0	0	0	1	0	2	0.50	-18.380	-17.541	-18.380
0	0	0	0	0	1	0	2	0.75	-21.640	-21.253	-23.241
0	0	0	0	0	1	0	2	1.10	-25.710	-28.681	-25.732
0	0	0	0	0	1	0	4	0.50	-7.010	-5.945	-12.447
0	0	0	0	0	1	0	4	0.75	-13.390	-17.001	-17.493
0	0	0	0	00	1	0	4	1.10	-31.720	-25.84	-34.927
0	0	0	0	0	1	0	6	0.50	-0.060	-3.047	-1.834
0	0	0	0	0	1	0	6	0.75	-21.030	-15.327	-21.525

0	0	0	0	0	1	0	6	1.10	-21.360	-25.325	-21.358
0	0	0	0	0	1	0	8	0.50	-5.040	-4.829	-9.514
0	0	0	0	0	1	0	8	0.75	-13.300	-15.331	-10.445
0	0	0	0	0	1	0	8	1.10	-31.860	-30.932	-34.709
0	0	0	0	0	1	0	10	0.50	-13.600	-12.811	-9.683
0	0	0	0	0	1	0	10	0.75	-22.040	-22.258	-11.752
0	0	0	0	0	1	0	10	1.10	-45.080	-45.668	-45.198
0	0	0	0	0	0	1	2	0.50	0.940	0.712	0.167
0	0	0	0	0	0	1	2	0.75	0.680	-1.299	0.167
0	0	0	0	0	0	1	2	1.10	0.530	0.995	0.167
0	0	0	0	0	0	1	4	0.50	-0.100	1.639	0.167
0	0	0	0	0	0	1	4	0.75	-0.050	0.192	0.167
0	0	0	0	0	0	1	4	1.10	-0.070	0.527	0.167
0	0	0	0	0	0	1	6	0.50	0.360	0.826	0.167
0	0	0	0	0	0	1	6	0.75	0.250	-0.891	0.167
0	0	0	0	0	0	1	6	1.10	0.530	-0.087	0.167
0	0	0	0	0	0	1	8	0.50	0.190	-0.426	0.167
0	0	0	0	0	0	1	8	0.75	0.340	0.699	0.167
0	0	0	0	0	0	1	8	1.10	0.470	2.320	0.167
0	0	0	0	0	0	1	10	0.50	-0.060	-0.403	0.167
0	0	0	0	0	0	1	10	0.75	-0.180	2.053	0.167
0	0	0	0	0	0	1	10	1.10	0.270	-0.366	0.167
1	0	0	0	0	0	0	5	0.75	-24.85	-23.752	-21.210
0	1	0	0	0	0	0	10	1.00	-77.02	-73.886	-72.729
0	0	1	0	0	0	0	6	0.60	-4.148	0.188	-1.706
0	0	0	1	0	0	0	3	0.75	-72.87	-67.586	-77.57
0	0	0	0	1	0	0	8	0.90	-94.49	-99.317	-94.004
0	0	0	0	0	1	0	7	1.10	-26.61	-26.901	-21.195
0	0	0	0	0	0	1	5	1.10	0.23	-0.339	0.167

Table 2. Precise comparison of different algorithm about reduction in yield of rice

Algorithm	Training times	Training sample total error	Training sample average error	Test sample total error
SABP	3213	309.201	2.945	19.538
AGABP	---	729.29	6.946	22.765

4 Conclusions

(1) From Table 2, it can be seen that the SABP algorithm is good on training time and forecasting precision. The new method can handle the disadvantage of the original BP algorithm. It is obvious that the rice inundated function in each growth period can be described by one ANN.

(2) Due to the application of SABP algorithm，it may form computer process. Thus it can extend conveniently to solve other water resources engineering in relate to analyzing and forecasting. So, it is significant practically to research the SABP application in water resources development.

Correspondence to:
Qiang Fu, Wei Zu
School of Water Conservancy & Civil Engineering
Northeast Agriculture University
Harbin, Heilongjiang 150030, China
fuqiang100@371.net

References

[1] Fu Q, Liang CH. Modeling and optimize technique of water saving irrigation system. Sichuan Technology Publishing Company, Chengdu, Sichuan, China. 2002;6:56-62.
[2] Fu Q, Song YF, Xiao JM. Forecast well irrigated paddy rice water demand based on artificial neural net. Northeast Hydraulic and Hydroelectric Publishing Company. 2002;5:38-40.
[3] Fu Q, Fu H. Applying PPE model based on RAGA in the investment decision-making of water saving irrigation project. Nature and Science 2003;1(1):57-61.
[4] Fu Q, Zu W. Using multi-dimension dynamic planning based on RAGA to optimize irrigation system under non-sufficient irrigation condition. The Journal of American Science 2005;1(1):68-74.
[5] Jin JL, Ding J. Heredity algorithm and application in water science.. Sichuan College Publishing Company. 2000:147-52.
[6] Kang LSH, Yun Xie. Simulated annealing algorithm. Technology Publishing Company. 2000:22.
[7] Ma XD, Fu Q. Applying self-organizing competition artificial neural networks to classify the soil. Nature and Science 2003;1(1):75-81.
[8] Zhu RL. Experiment research of paddy rice inundated grade influent on production. Downfold disaster and fathering experiment. Dalian Publishing Company. 1990:146-56.
[9] Zhou JL. Heredity algorithm and application in water science. Sichuan University Doctor Dissertation. 2000:93-5.

Vaccination against Schistosoma *mansoni* Infection by DNA Encoding SM 21.7 Antigen

Mahmoud H. Romeih[1*], Ahmed M. Hanem[1**], Tarek S. Abou Shousha T.S[2], Mohamed A. Saber[1]

Biochemistry and Molecular Biology [1] and Pathology [2] Departments, Theodor Bilharz Research Institute, Giza, Egypt
Present address: * Biochemistry and Molecular Biology Department, Michigan State University, USA; ** Food Science and Human Nutrition, Michigan State University, USA
romeih@msu.edu; ahmedha@msu.edu

Abstract: The current focus of schistosomiasis research is to develop a vaccine that will significantly reduce the incidence of disease. Immunization with DNA is a new trend in vaccine development that could enhance the safety and efficacy of currently used vaccines. The immunogenicity and protective efficacy of a DNA vaccine encoding the antigen SM21.7 was evaluated in C57 BL/6 and Swiss albino mice. The ORF of SM21.7 has been cloned into the eukaryotic expression vector pcDNA 1/Amp under control of CMV late enhancer promoter. The groups of mice were vaccinated intramuscularly with SM21.7-pcDNA1 and boosted twice shown high and specific humoral response in comparison with control (blank pcDNA1/Amp). The level of anti SM21.7 antibody in immunized mice at weeks 3, 6 and 9 intervals post-immunization was significantly higher than in control group. The maximum level of antibodies was obtained at 7 weeks post - challenge infection. Sera from immunized mice could recognize the SM21.7 protein and the specific antibodies were able to mediate a significant killing of schistosomula using peritoneal macrophages as effector cells. In contrast the preimmune sera and sera control serum had no specific reactivity to SM21.7 protein. Immunization with SM21.7- pcDNA conferred a significant level of protection against challenge (41, 53%) in Swiss Albino and C57BL/6 mice respectively. Histopathological examination of the vaccinated liver revealed a decreased in the number, size and change in the cellular of the granuloma compared to the control infected liver. In addition reductions in worm viability, worm fecundity and egg hatching ability have been observed following challenge with *Schistosoma mansoni* cercariae. The number of eggs in the liver and intestine was reduced by 62% and 67% respectively compared to control group. The results suggested that SM21.7 might be a candidate antigen for the generation of antipathology vaccine against schistosomes. [Nature and Science. 2005;3(2):28-35].

Keywords: *Schistosoma mansoni*; DNA vaccine, pcDNA1/Amp, 21.7kDa schistosomula protein; DNA immunization; and vaccine

1. Introduction

Schistosomiasis caused by blood flukes of the genus *Schistosoma* is the second most important parasitic disease of humans. The World Health Organization has estimated that 200 million people are infected worldwide in 74 countries and a further 600 million are at risk of this disease (Chisel *et al*., 2000; WHO, 2002). In terms of control strategies for schistosomiasis, enormous efforts have been made to develop effective anti-schistosome vaccines, but none is available at the present time (Shuxian *et al*., 1998). The complexity of the schistosome and the life cycle of this parasite may, at least partly, contribute to the difficulties associated with vaccine development. Pathology and morbidity in this chronic disease result from the host immune inflammatory response to parasite eggs trapped in the liver and other organs (Berqiust, 1998; Capron *et al*., 2001).

The development of the drug praziquantel (Davis and Wegner, 1979) and what proved to be of equal importance, its drastic reduction in price, rapidly made chemotherapy the cornerstone of control (WHO 1985,1993) leading to a dramatic reduction of morbidity in endemic areas (WHO, 1985; WHO, 2001). However, chemotherapy does not affect transmission of the infection and high re-infection rates, even after mass treatment, limit the success by demanding frequently re-scheduled treatments. For this reason, a complimentary approach that can be integrated with chemotherapy, i.e. vaccination, is very much needed (Bergquist 1992).

Vaccination can either be targeted towards the prevention of infection or to the reduction of parasite fecundity. A reduction in worm numbers is the "gold standard" for anti-schistosome vaccine development but, as schistosome eggs are responsible for both pathology and transmission, a vaccine targeted on parasite fecundity and egg viability also appears to be entirely relevant (Taylor 2002).

Recently, a significant effort has been made to develop a protective vaccine against schistosome infections, and several vaccine candidate have been identified (Ahmed *et al*., 2001; Da'dara *et al*., 2001; Zhang *et al*., 2001; Dupre *et al*., 2001; Capron *et al*., 2001; McManus, 2000; Waine and McManus, 1997; Shuxian *et al.,* 1998), but as the efficacy of any of these against schistosomiasis remains uncertain, the identification and characterization of new anti-schistosome vaccine molecules remains a priority. The

development of vaccine remains to be an important long-term and remain a challenging goal in the control of schistosomiasis (Hafalla *et al*., 1999).

DNA vaccine is an attractive and novel immunization strategy against a wide range of infectious diseases and tumers. Injection of plasmid DNA as vaccine was first demonstrated to be effective using influenza as model, where it was shown that DNA encoding nucleoprotein (NP) induced cytotoxic T-lymphocytes (CTLs) and cross-strain protection of mice (Ulmer *et al*., 1993). The effectiveness of DNA vaccines against viruses, parasites, and cancer cells has been demonstrated in animal models (Tuteja, 1999). It has been shown that DNA immunization induces both antigens - specific cellular and humoral immune response (Ramsay *et al*., 1999; Alarcon *et al*., 1999; Gurnathan *et al*., 2000; Kowalaczyk, 1999). Nucleic acid vaccination against schistosomiasis has lately been investigated using a panel of plasmid encoding schistosome antigenic proteins such as Sjc 26 GST, Sj79 (Zhou *et al.* 1999a*, Zhang et al,* 2001*), Schsitosoma japonicum* paramyosine (Yang *et al.,* 1995; Zhou *et al.,* 1999b), and *Schsitosoma m* 23, 28 GST from *Schsitosoma mansoni* (Dupre *et al*., 1997).

One of the major goals in our laboratory is to develop a protective vaccine against *Schsitosoma mansoni* infection. Intramuscular immunization with SM21.7 protein induced a strong immune response that resulted in a significant reduction in the number of the adult worm (40, 71%) in Swiss Albino and C57BL/6 mice respectively (Ahmed *et al*., 2001). In this study to evaluate the immunogenicity and protective efficacy of SM21.7 as a DNA vaccine. The ORF of SM21.7 *Schsitosoma mansoni* antigen was clones into the pcDNA 1/Amp vector. Vaccination with the naked DNA encoding SM21.7 induced significant levels of specific IgG antibody responses and it has been successful in inducing significant levels of protection in two strains of mice as being assessed histopathologically and parasitologically.

2. Materials and Methods

2.1 Parasites and animals

An Egyptian strain of *Schistosoma mansoni* cercariae and naive female Swiss albino mice and C57 BL/6 (6 – 8 weeks of age), were obtained from the Schistosome Biological Supply Program, Theodor Bilharz Research Institute, Giza, Egypt. The animals were separated in groups of twenty and they kept under standard laboratory conditions and used for vaccine trials.

2.2 Construction and preparation of SM21.7 plasmid DNA

The SM21.7-pcDNA plasmid encoding the full length SM21.7 was used throughout these experiments. The cDNA containing the entire coding region of SM21.7, which has been isolated from schistosomula cDNA library (Ahmed *et al*., 2001). The ORF was excised from the recombinant pBluescript vector using polymerase chain reaction (Saiki *et al*., 1988). A pair of primers was synthesized according to the DNA sequence of the SM21.7, *Bam*H1 adaptors linked to forward and reverse primers and the Kozark sequence was added to the position of initiator. The forward primer was 5′CATCTGGATCCATGGATAGTCC and the reverse 5′ TAACGGATCCCTAGTTACTTGG. The amplified sequence was ligated into the eukaryotic expression pcDNA1/Amp expression vector (Invitrogen, Corp, SanDiago, CA), which was previously digested with *Bam*H1 and treated with alkaline phosphates. The structure was verified by restriction digestion and sequencing. Large-scale preparation of the plasmid was carried out by using the alkali lysis method, followed by double banding on CsCl-EtBr gradient (Sambrook *et al.,* 1989). Then DNA was resuspended in phosphate buffer saline (PBS) for vaccination.

2.3 DNA vaccination

The groups of female Swiss albino and C57BL/6 mice were vaccinated by intramuscular injection. Two groups of each strain were used, three weeks after the first injection the groups of mice were subsequently boosted with 100 μg/ml of DNA, and after another 3 weeks, mice were boosted a second time with 50 μg/ml of DNA (He *et al*., 1997). The mice were challenged with 100 *Schsitosoma mansoni* cercariae by tail immersion method and perfused 7 weeks later (Duvall and DeWitt, 1967). Blood samples were collected from tail veins of all mice prior to immunization and thereafter at 3, 6 and 9 weeks intervals, as well as finally at 7 weeks post-challenge. Pooled serum samples were prepared from each group by mixing an equal volume of serum from each group, then used for ELISA and western blots analysis.

2.4 Antibody assay

To determine, the presence of SM21.7 antibody titer in collected sera from mice vaccinated with SM21.7-pcDNA1 and control (pcDNA1/amp alone). The preimmune and post vaccination sera were tested for specific immunoglobulin G (IgG) by an enzyme linked immunosorbant assay (ELISA) and Western blot analysis. The antigen used in ELISA was SM21.7 KDa purified protein, expressed in pET-3a expression vector and partially purified as described previously by Ahmed *et al*., (2001). SWAP was prepared to detect the specific antibodies of SM21.7-pcDNA1 by Western blot analysis.

ELISA was carried out in micrometer plates coated with purified SM21.7 antigen (3μg/well) prepared in PBS. The protocol of Zhang *et al*. (1998) was followed to detect the SM21.7 specific antibodies. The secondary antibody used in the ELISA was alkaline phosphatase-conjugated goat anti-mouse immunoglobulin G (Serotec Ltd, England).

2.5 Western blot analysis

Western blot analysis was performed as described by Yang *et al*., (1995). Briefly, SWAP was separated by 8% SDS-polyacrylamide gel (Laemmli, 1970), and transferred to a nitrocellulose membrane using a Bio-Rad protein transfer unit (Mini-gel). The nitrocellulose membranes were blocked in TBST (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.05% Tween 20) containing 3% bovine serum albumin (BSA) for 3 hours. Blots were probed with sera from mice immunized with the SM21.7-pcDNA1 parallel to pcDNA1/Amp as a control for 2 3 hours with gentle shaking at RT. Then blots were washed four times in TBST and incubated with alkaline phosphatase conjugated anti-serum for 1 hour. After incubation, the blots were washed in TBST and soaked in alkaline phosphatase substrate solution 5-bromo-4-chloro-3-indolyl-1-phosphate (BCIP) and nitroblue tetrazolium (NBT).

2.6 Infection and determination of worm burden

Three weeks after final boosting vaccinated and control mice groups were challenged with 100 *Schsitosoma mansoni* cercariae by tail immersion (Olivier and Stirwalt, 1952). The percentage of the resistance was calculated by perfusion of the adult worms from the portal veins at 7 weeks after challenge infection. The worm reduction rate (% protection) was calculated and the livers, spleen and intestines were collected and the eggs were counted according to the method described by Shuxian *et al.* (1997).

2.7 Enumeration of eggs in liver and intestine

Whole livers and intestines, of vaccinated and control mice were weighed and a known portion (0.5 g) was removed to a screw cap glass tube and frozen until digestion. For digestion 5ml of 5% KOH (potassium hydroxide) was added to each tube and incubated at 37 °C until the tissue was completely digested (10-12 hr). A 50 µl of the digest was placed on a glass slide and eggs were counted under a microscope (Liu *et al*., 1995). Total egg counts were expressed for each group of mice as the mean number of eggs /gram of mouse liver or intestine.

2.8 Histopathlogical examination

After scarification of different groups of animals, part of the liver tissue was immediately fixed in 10% buffered formalin solution and processed in paraffin blocks. 5µm thick sections were cut on albuminized glass slides and stained by Hematoxyline and eosine for routine histopathological examination, granuloma count, measurement of granuloma diameter, cellular profile and Masson trichrome staining for assessment of tissue fibrosis.

Liver egg-granulomas were counted in 5 successive low power fields (10X), and their diameters were measured using graduated eyepiece lens, considering only lobular granulomas containing central ova. Two perpendicular maximal diameters were measured, getting the mean diameter for each granuloma, and then calculating the mean granuloma diameter for the group.

The cellular profiles of liver egg-granulomas were studied, with calculation of the percentage of different types of cells forming the granuloma in the different groups. The type of the granuloma, whether cellular, fibrocellular or fibrous was defined according to the cellular to fibrous ratio. The percentage of egg-granulomas containing intact or degenerated miracidia was also calculated.

2.9 Statistic

Statistical significance was determined by student's t-test and significance was determined using a p value < 0.05 as being significant.

3. Results

3.1 Antibody levels

The intramuscular injection with pcDNA1-SM21.7 stimulated high titer of antibody responses in Swiss albino mice and C57 Bl/6 mice. The sera collected from mice at 3 weeks post immunization were able to detect the purified SM21.7 protein with highly titer and there was no significant difference in the titer at 6 or 9 weeks post immunization. The maximum peak occurred at 7 weeks post challenge infection in the immuno-challenge with SM21.7-pcDNA1 groups (Figure 1). In contrast, the preimmune sera and sera from the pcDNA1/Amp blank vector immunized mice had no specific reactivity to SM21.7 protein. Western blots showed sera from pcDNA1-SM21.7 immunized mice could recognize native SM21.7 from SWAP (Figure 2).

3.2 Protection induced by DNA vaccine

To determine whether the SM21.7-pcDNA1 vaccine conferred protective immunity, and to determine whether the most of DNA immunization affect immunogenicity and/or vaccine efficacy. Vaccinated mice and control were challenged with 100 cercariae each and the number of worms recovered seven weeks later was assessed. The egg counts showed a highly significant difference either in terms of reduced worm burden or egg number present, isolated from the liver or between mesenteric. In the three separate experiments, SM21.7-pcDNA1, inoculated intramuscularly three times at doses of 100 or 50 µg per/mouse, was found to provide significant worm reduction rates ($P < 0.001$). The animals showed a significant level of protection (41, 53%) in Swiss albino and C57BL/6 mice respectively following challenge infection (Figure 3). There was a subsequent reduction in the number of eggs present in the livers and intestine of the pcDNA1-SM21.7 immunized group when compared with the pcDNA1/Amp blank vector as a control group.

3.3 Effect on egg count in liver and intestine

There was a substantial reduction in the mean number of eggs/gm tissue of liver and intestine of the

immuno-challenge group mice (p∠0.001), when compared to controls (Figure 4). The number of eggs in the liver and intestine was reduced by 62% and 67% respectively compared to control group.

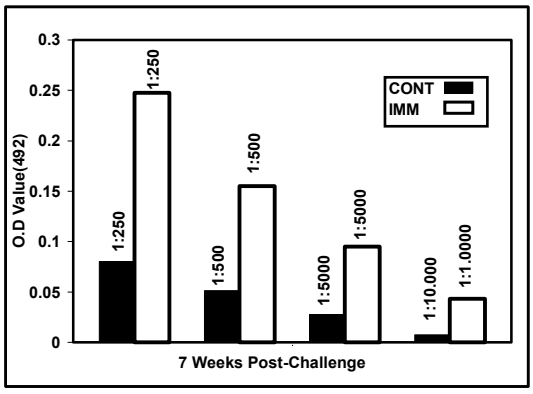


Figure 1. ELIA analysis of control and immunized groups at 7 weeks post-challenge. ELISA titer of specific IgG of tail blood C57 BL/6 (n=15) vaccinated with pcDNA1 /SM21.7 (n=15) at 7 week post challenge. Sera were diluted in TBS containing 0.05% Tween 20 [1:250 (p< 0.0001); 1:500 (p< 0.001); 1:5000 (p< 0.01); 1:10,000 (p < 0.01)]. Note. [CONT=control; IMM = immunized). The control group was vaccinated with pcDNA1/Amp alone. The data are representative of three successive experiments. We have obtained approximately the same results with slightly differences of Swiss albino mice.

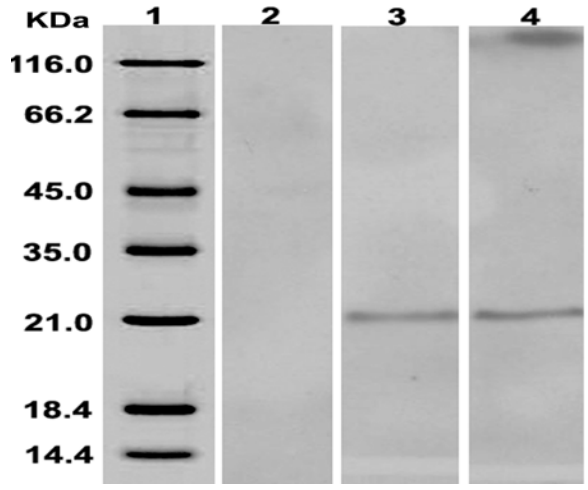
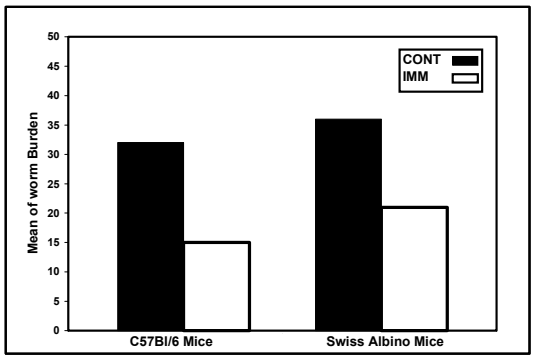
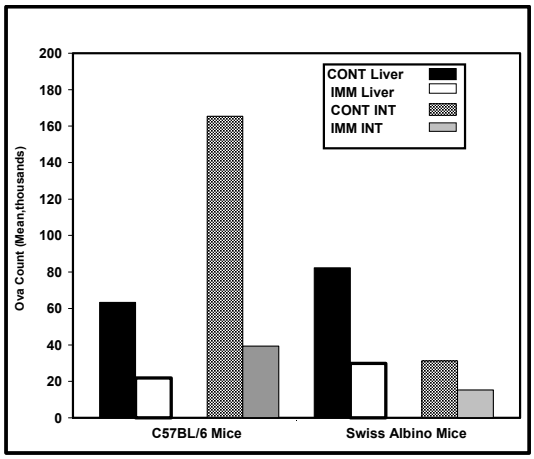


Figure 2. Western blot analysis showing the sera from mice immunized with pcDNA1 blank vector had no specific reactivity to SM21.7 protein (lane 1). Sera from mice immunized with SM21.7-pcDNA recognize native SM21.7 from SWAP on Black/6 and Swiss albino mice (Lane 3, 4). Lane (1) high molecular weight standards.



Worm Burden (Mean) in Immunized and Control Mice

Figure 3. Changes in worm burden % in immunized and control liver histopathology sections. Mice were vaccinated with pcDNA1/ SM21.7 and the control group was vaccinated with pcDNA1/Amp alone. The percentage of protection was calculated by perfusions of adult worms from the portal vein at 7 week post-challenge infection. The percentage of protection ranged from (41, 53%) in Swiss albino mice and C57 Bl/6 mice respectively (p<0.0001). Note [CONT=control and IMM = immunized] and the control groups were vaccinated with pcDNA1 1/Amp alone.



Changes in Ova Count (Mean eggs) Immunized Liver and Intestine Vessels and Control Mice

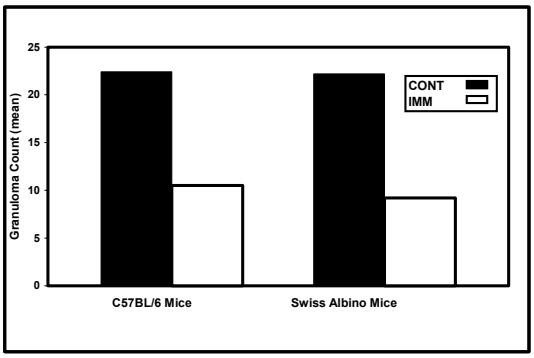
Figure 4. Changes in ova count in control and immunized livers and mesenteric. The percentage of degenerated ova in control group is much higher than immunized group in the two strains of mice (Swiss albino mice and C57 BL/6). There is highly significant difference from both control and immunized group P<0.0001. Note [L=liver and INT=intestine], and the control groups were immunized with pcDNA1 1/Amp alone.

3.4 Histopathological changes in granuloma

Liver sections of both immunized and control groups at 7 weeks post infection were studied for granuloma count and size. The histopathological examination showed a significantly greater number of egg granulomas in control group than in immunized group (Figure 5). The mean diameter of granuloma was significantly higher in control group compared to the immunized group (p∠0.001) as shown in (Figure 6). In addition the percentages of degenerated ova were higher in the immunized group compared to the control groups.

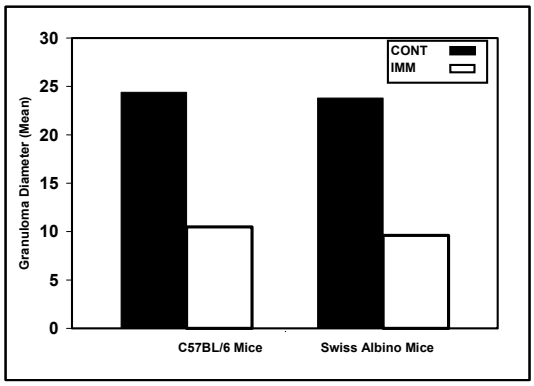
Sections of infected mice liver with *S. mansoni* cercariae and immunized with SM21.7-pcDNA1 showed, lesser number of smaller egg granuloma usually formed of central egg surrounded by lymphocytes, epithelioid histiocytes, fibroblasts and fewer peripherally located eosinophils and neutrophils. In contrast the control revealed greater number of larger egg granulomas formed of an ovum surrounded by large number of eosinophilis and neutrophils as well as some macrophages and focal area of eosinophilic necrosis (Figure 7). On the other hand, Masson's trichrome staining showed more fribrocellular granuloma formed of central ova surrounded by more macrophages, lymphocytes, fibroblast and collagen fibers in the vaccinated group compared to the control (Figure 8).

The most striking feature regard the cellular profile of egg granulomas was the percentage of eosinophils and neutrophils was much greater in the granulomas of the control group than immunized group. The percentage lymphocytes and macrophages were greater in the immunized group than the control group. The percentage of ova containing degenerated miracidia was also greater in the immunized group compared to the control group.



Changes in Granuloma Count in Immunized and Control Mice

Figure 5. Changes in granuloma count between vaccinated and control infected. There was a remarkable decrease in the granuloma count in immunized group rather than control group. Significant difference between control and immunized is (P<0.001). Note [CONT=control, IMM=immunized] and the control group was immunized with pcDNA1/Amp alone.



Changes in Granuloma Diameter in Immunized and Control Mice

Figure 6. Changes in granuloma diameter (means 5 in microns) in immunized and control mice. There is a highly significant differences in immunized compared with control group. The mean diameter in control group is higher than immunized group (p<0.001). Note. [CONT=control and IMM=immunized], and the control group was vaccinated with pcDNA1 1/Amp alone.

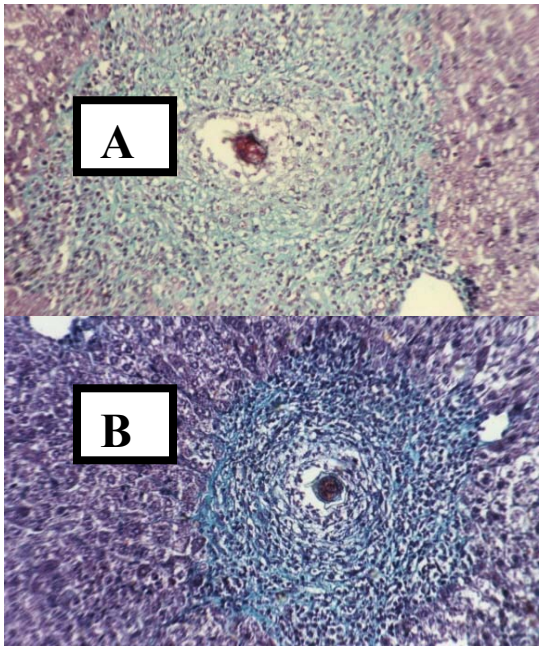


Figure 7. Histopathological examination of vaccinated control infected mice using Masson's stain. A: Section in mouse liver infected with *S. mansoni* showing a large egg-granuloma, formed of central ova surrounded by inflammatory cells and irregularly deposited collagen fibers (Masson's trichrome stain, X200). B: Section in mouse liver infected by *S. mansoni* cercariae and vaccinated by using pcDNA/SM21.7, showing smaller fibrocellular granuloma formed of central ova, surrounded by histiocytes, lymphocytes, fibroblasts and concentric collagen fibers (Masson's trichrome stain, X200).

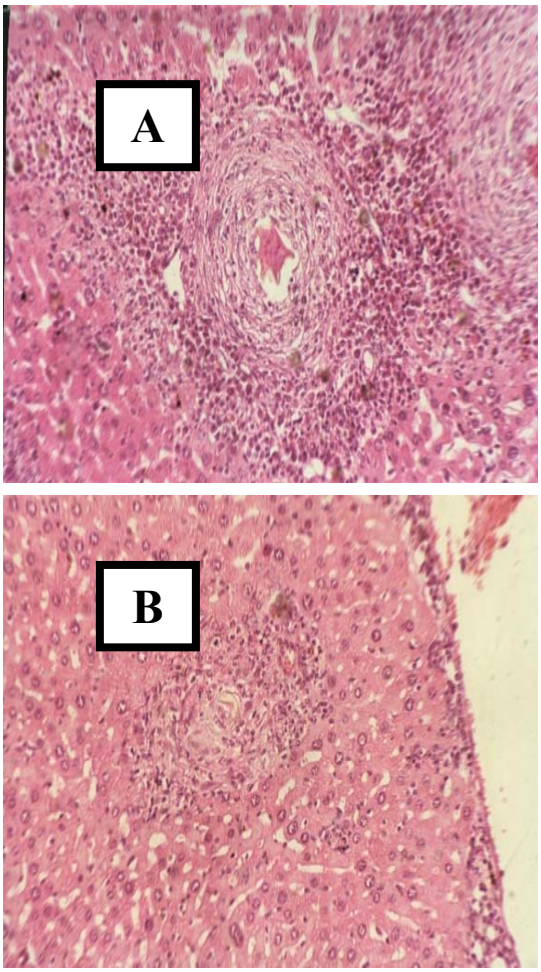


Figure 8. Histopathological examination of vaccinated and control infected mice using Haematoxylin and Eosin stain A: Section in mouse liver infected with *S. mansoni* showing a large portal egg-granuloma, formed of central ova surrounded by large number of eosinophils, neutrophils and histiocytes (Haematoxylin and Eosin stain, X200). B: Section in mouse liver infected by *S. mansoni* cercariae and vaccinated by using pcDNA/SM21.7, showing smaller egg granuloma formed around remnant of ova in the center, surrounded by some mononuclear inflammatory cells and few eosinophils (Haematoxylin and Eosin stain, X200).

4. Discussion

Nucleic acid immunization can be an effective vaccination technology that delivers DNA constructs encoding specific immunogens into host cells, inducing both antigen-specific humoral and cellular immune responses (Bergquist, 2002). Since the first demonstration of protective immunity against viral challenge induced by DNA vaccination using a plasmid DNA encoding influenza A nucleoprotein. Many trials with various degrees of success have been achieved (Ulmer *et al*., 1993), and the main methods of plasmid-DNA application are intramuscular injection and intradermal delivery into skin (Smahel, 2002). In the case of schistosomiasis, vaccination with DNA has been shows to induced immune responses in rats (Capron *et al*., 1997; Dupre *et al.,* 1997), and partial protection against challenge in mice (Mohamed *el al.,* 1998) underlining the potential of this method of vaccine delivery for this disease.

Previously, we developed a candidate vaccine against *S. mansoni* infection based on the SM21.7 tegument protein, isolated from an Egyptian *S. mansoni* strain liver worms cDNA library by immunoscreening using vaccinated rabbit sera. Intramuscular immunization with SM21.7 protein induced a strong immune response that resulted in a significant reduction in the number of adult worm 40 and 71% in Swiss Albino mice and C57BL/6 mice (Ahmed *et al*., 2001). In this study to evaluate the immunogenicity and protective efficacy of SM21.7 as a DNA vaccine. The ORF of SM21.7 *Schsitosoma mansoni* antigen was clones into the pcDNA 1/Amp vector. Vaccination with the naked DNA encoding SM21.7 induced significant levels of specific IgG antibody responses and it has been successful in inducing significant levels of protection in two strains of mice as being assessed histopathologically and parasitologically.

In this study the data are representative of three successive experiments and attempts were made in the current studies to maximize the level of response to the injected DNA. Firstly; Swiss albino and C57 BL/6 mice were chosen because these strains previously demonstrated a high level of protection against *Schsitosoma mansoni* using SM21.7 recombinant protein (Ahmed *et al*., 2001). The immunization with reSjc26GST in C57BL/6 mice elicited higher humoral and cellular immune responses in C57BL/6 mice than the BALB/c mice (Shuxian *et al*., 1998). Secondly, the direct intramuscular injection of plasmid DNA has been used to induce immune responses, and it has been found that a simple saline solution appears to be a suitable carrier resulting in transfection of between 1-5% of myofibrils in the vicinity of the injection site in the case of intramuscular administration (Wolff *et al*., 1990). Thirdly, the pcDNA1/Amp was chosen because it has been used by many of investigators successfully to obtain partial levels of protections (Dupre et *al*., 1997; Mohamed *et al*., 1998; Zhou *et al.,* 1999).

To determine if the SM21.7-pcDNA vaccine conferred protection against *Schistosoma mansoni*, all groups were challenged with 100 cercariae 3 weeks after the last boost and, 7 weeks later, worm burdens were analyzed. In all cases, the SM21.7-pcDNA vaccine induced statistically significant levels of protection to *Schsitosoma mansoni* cercarial challenge infection. The administration of SM21.7-pcDNA1 by intramuscular inoculation in this study resulted in expression *in vivo*, which induced specific antibodies in Swiss albino and Black/6 mice, which could be detected by Western blot analysis (Figure 2) and ELISA (Figure 1).

The reduction in worm burden in animals immunized with SM21.7-pcDNA was significantly higher than in animals immunized with control. Our results demonstrated that the level of pcDNA-SM21.7 antibodies in immunized is significantly higher than the control group of mice ($p< 0.001$), at 6 and 9 weeks post immunization. The maximum peak was recognized at 7 weeks post-challenge infection in the two strains of mice (Figure 1). In addition vaccination of Swiss albino and C57 BL/6 mice conferred a significant level of protection (41 and 53%) in Swiss albino mice and C57 Black 6/mice against challenge infection (Figure 3).

The protection in the present study is much higher that obtained by Yu *et al.* (2002) with SjC 21.7-pcDNA3 in BALB/C mice. They reported the worm reduction rate was 29.9% and its egg reduction rate 13.8% in the test group; 31.9% and 28.0% respectively in the boost group. The egg reduction rate in the boost group was higher than that of the test group ($P < 0.05$). On the other hand, Hafalla *et al*. (1999), have been reported that, *Schsitosoma japonicum* molecule, Sj21.7, is a target of IgE antibodies from high-IgE/SWAP responders, indicating that it may be an important vaccine candidate against human schistosomiasis japonica.

Vaccination can be targeted towards the prevention of infection or to the reduction of parasite fecundity. A reduction in worm numbers is the "gold standard" for anti-schistosome vaccine development but, as schistosome eggs are responsible for both pathology and transmission, a vaccine targeted on parasite fecundity and egg viability also appears to be entirely relevant (Capron *et al*., 2002). The effective vaccine would prevent the initial infection and might reduce egg granuloma associated pathology (McManus *et al*., 1998). The histopathological examination in the present study of the liver revealed a decreased in the number and size of egg-granulomas in the liver, and intestine of vaccinated mice compared with blank vector injected mice. These findings pointed to significant reduction of parasite fecundity and egg viability, the latter directly affecting transmission potentialities of the disease.

Our results are in agreement with, Hassanein *et al* (1997), they have been reported that the reduction in granuloma number and size was associated with amelioration of pathological changes in SEA immunized group. The results present in this study are very close to the result of (Zhou et al., 199a), they have been reported, reduction in worm burden following exposure to infection or reinfection and reduction of pathology by a decrease in worm fecundity, directly affecting the transmission of *S. japonicum C26GST* recombinant protein.

5. Conclusion

We have shown that the intramuscular immunization with pcDNA /SM21.7 was able to induce the immune responses in C57BL/6 and Swiss albino mice. Sera from immunized mice could recognize the SM21.7 protein and the specific antibodies were able to mediate significant macrophages as effecter cells. In addition, our results suggested that the vaccination was able to confer a protective immunity efficacy in C57 BL/6 mice. In addition the pcDNA/SM21.7 showed potential as a DNA vaccine and anti-pathological vaccine. Further studies should yield more insight on the vaccine potential of this antigen and its mechanism of action.

Acknowledgement

We thank Dr. Joseph F. Leykam, Director of Genomics Technology Support Facility, Department of Biochemistry and Molecular Biology, Michigan State University for critical review of the manuscript.

Correspondence to:
Mahmoud Romeih
Associate Prof of Biochemistry and Molecular Biology
Biochemistry and Molecular Biology Department
Theodor Bilharz Research Institute, Giza, Egypt.
Email: romeih@msu.edu, romeihm@yahoo.com

5. References

[1] Ahmed HM, Romeih MH, Sheriff SA, Fathom FA, Saber MA. Protection against *Schistosoma mansoni* infection with recombinant schistosomula 21.7 kDa protein. Arab Journal of Biotechnology 2001;24:229-49.

[2] Alarcon JB, Waine GW, McManus DP. DNA vaccines: technology and application as anti-parasite and anti-microbial agents. Adv Parasitoly 1999;42:343-410.

[3] Bergquist NR. Schistosomiasis research funding: the TDR contribution. Mem Inst Oswaldo Cruz 1992;IV(87 Suppl):153–61.

[4] Bergquist NR. Schistosomiasis vaccine development: approaches and prospects. Mem Inst Oswaldo Cruz 1995;90(2):221-7.

[5] Bergquist NR. Schistosomiasis: from risk assessment to control. Trends Parasitol. 2002:18(7):309-14.

[6] Capron M, Torpier G, Capron A. In vitro killing of *S. mansoni* schistosomula by eosinophils from infected rats: role of cytophilic antibodies. J Immunol 1979:123(5):2220-30.

[7] Capron A, Riveau GJ, Bartley PB, McManus DP. Prospects for a schistosome vaccine. Curr Drug Targets Immune Endocr Metabol Disord 2002;2(3):281-90.

[8] Chisel L, Engels D, Montresor A. Savioli L. The global status of schistosomiasis and its control. Acta Trop 2000;77:41–51

[9] Da'dara AA, Skelly PJ, Wang MM, Harn D. AImmunization with plasmid DNA encoding the integral membrane protein, Sm23, elicits a protective immune response against schistosome infection in mice. Vaccine 2001;:20:359-69.

[10] Davis A, Wegner DH. Multicentre trials of praziquantel in human schistosomiasis: design and techniques. Bull World Health Organ 1979;57:767-71.

[11] Dupre L, Poulain-Godefroy O, Ban E, Ivanoff N, Mekranfar M, Schacht AM, Capron A, Riveau G. Intradermal immunization of rats with plasmid DNA encoding *Schistosoma mansoni* 28 kDa glutathione S-transferase. Parasite Immunol 1997;19(11):505-13.

[12] Dupre L, Kremer L, Wolowczuk I, Riveau G, Capron A, Locht C. Immunostimulatory effect of IL-18-encoding plasmid in DNA vaccination against murine *Schistosoma mansoni* infection. Vaccine 2001;19:1373-80.

[13] Duvall RH, DeWitt WB. An improved perfusion technique for recovering adult schistosomes from laboratory animals. Am J Trop Med Hyg 1967;16(4):483-6.

[14] Gurunathan S, Wu CY, Freidag BL, Seder RA. DNA vaccines: a key for inducing long-term cellular immunity. Curr Opin Immunol 2000;12:442-7.

[15] Hafalla JC, Alamares JG, Acosta LP, Dunne DW, Ramirez BL, Santiago ML. Molecular identification of a 21.7 kDa *Schistosoma japonicum* antigen as a target of the human IgE response. Mol Biochem Parasitol 1999;98(1):157-61

[16] He J, Hoffman SL, Hayes CG. DNA inoculation with a plasmid vector carrying the hepatitis E virus structural protein gene induces immune response in mice. Vaccine 1997;15(4):357-62.

[17] Hassanein H, Akl M, Shaker Z, El-Baz H, Sharmy R, Rabiae I, Botros S. Induction of hepatic egg granuloma hypo responsiveness in murine schistosomiasis mansoni by intravenous injection of small doses of soluble egg antigen. APMIS 1997;105(10):773-83.

[18] Hota-Mitchell S, Clarke MW, Podesta RB, Dekaban GA. Recombinant vaccinia viruses and gene gun vectors expressing the large subunit of *Schistosoma mansoni* calpain used in a murine immunization-challenge model. Vaccine 1999;17(11-12):1338-54.

[19] Kowalczyk DW, Ertl HC. Immune responses to DNA vaccines. Cell Mol Life Sci 1999;55:751-70.

[20] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;15:680.

[21] Liu S, Song G, Xu Y, Yang W, McManus DP. Immunization of mice with recombinant Sjc26GST induces a pronounced anti-fecundity effect after experimental infection with Chinese *Schistosoma japonicum*. Vaccine 1995;13(6):603-7.

[22] McManus DP, Liu S, Song G, Xu Y, Wong JM. The vaccine efficacy of native paramyosin (Sj-97) against Chinese *Schistosoma japonicum*. Int J Parasitol 1998;28(11):1739-42.

[23] McManus DP. DNA vaccine against Asian schistosomiasis: the story unfolds. Int J Parasitol 2000;30:265-71

[24] Mohamed MM, Shalaby KA, LoVerde PT, Karim AM. Characterization of Sm20.8, a member of a family of schistosome tegumental antigens. Mol Biochem Parasitol 1998; 96(1-2):15-25.

[25] Oliver L, Stirwalt MA. An efficient method for exposure of mice to cercariae *S. mansoni.* J Parasitology 1952;38:19-35.

[26] Ramsay AJ, Kent SJ, Strugnell RA, Suhrbier A, Thomson SA, Ramshaw I.A. Genetic vaccination strategies for enhanced cellular, humoral and mucosal immunity. Immunol Rev 1999; 171:27-44.

[27] Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 1988;239(4839):487-91.

[28] Sambrook J, Fritsch EF, Maniatis T. in Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY, 1989; Vol. (1, 2, 3).

[29] Shuxian L, Yongkang H, Gung Chen S, Xing-Song L, Yuxin X, McManus DP. Anti-fecundity immunity to Schistosoma japonicum induced in Chinese water buffaloes (Bos buffelus) after vaccination with recombinant 26 kDa glutathione-S-transferase (reSjc26GST). Vet Parasitol. 1997;69(1-2): 39-47

[30] Shuxian L., Guangchen S., Yuxin X., McManus D.P, Hotez P.J. Progress in the development of a vaccine against schistosomiasis in China. Int J Infect Dis 1998:2(3):176-80.

[31] Smahel M. DNA Vaccine. Cas Lek Cesk 2002;22(141):26-32.

[32] Tuteja R. DNA vaccines: a ray of hope. Crit Rev Biochem Mol Biol 1999;34(1):1-24.

[33] Ulmer JB, Donnelly JJ, Parker SE, Rhodes GH, Felgner PL, Dwarki VJ, Gromkowski SH, Deck RR, DeWitt CM, Friedman A, *et al*. Heterologous protection against influenza by injection of DNA encoding a viral protein. Science 1993;259(5102) :1745-9.

[34] Waine GJ, Yang W, Scott JC, McManus DP, Kalinna BH. DNA-based vaccination using Schistosoma japonicum (Asian blood-fluke) genes. Vaccine 1997;15(8):846-8.

[35] World Health Organization (WHO). The Control of Schistosomiasis. Report of a WHO Expert Committee. WHO Technical Report Series 1985;728:113, Geneva.

[36] World Health Organization (WHO). The Control of Schistosomiasis. Second report of the WHO Expert Committee. WHO Technical Report Series 1993;830:86, Geneva.

[37] World Health Organization (WHO). The Prevention and Control of Schistosomiasis. and Soil-transmitted Helminthiasis. Report of the Joint WHO Expert Committees. 2002: WHO Technical Report Series (manuscript).

[38] Wynn TA, Cheever AW, Williams ME, Hieny S, Caspar P, Kuhn R, Muller W, Sher A. IL-10 regulates liver pathology in acute murine Schistosomiasis mansoni but is not required for immune down-modulation of chronic disease. J Immunol 1998;160(9):4473-80.

[39] Wolff J.A., Malone R.W., Williams P., Chong W., Acsadi G., Jani A., Felgner P.L. Direct gene transfer into mouse muscle in vivo. Science 1990; 247(4949 Pt 1):1465-8.

[40] Yang W, Waine G.J., McManus D.P. Antibodies to *Schistosoma japonicum* (Asian bloodfluke) paramyosin induced by nucleic acid vaccination. Biochem Biophys Res Commun. 1995;212(3): 1029-39.

[41] Yu CX, Zhu YC, Yin XR, Ren JG, Si J, Xu YL, Shen LN. Protective immunity induced by the nucleic acid vaccine of Sjc 21.7 in mice. Article in Chinese Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2002;20(4):201-4.

[42] Zhang Y, Taylor MG, Bickle QD. *Schistosoma japonicum* myosin: cloning, expression and vaccination studies with the homologue of the *S. mansoni* myosin fragment IrV-5. Parasite Immunol 1998;20(12):583-94.

[43] Zhang Y., Taylor M.G., Johansen M.V., Bickle Q.D. Vaccination of mice with a cocktail DNA vaccine induces a Th1-type immune response and partial protection against *Schistosoma japonicum* infection.Vaccine 2001;207(5-6):24-730.

[44] Zhou S, Liu S, Song G, Xu Y. Studies on the features of protective immune response induced by recombinant Sjc26GST of *Schistosoma japonicum*. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 1999a;17 (2):74-7.

[45] Zhou S, Liu S, Song G, Xu Y. Cloning, sequencing and expression of the full-length gene encoding paramyosin of Schistosoma japonicum in vivo. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 1999b;17(4):196-9.

The Optimization Plan of Urban Drainage System of Shenyang City

Jing Zhang, Junshi He, Jing Tian
College of Water Resources, Shenyang Agricultural University, Shenyang, Liaoning 110161, China;
Telephone: 024-88487134; E-mail: hejunshi@163.com

Abstract: This article takes the characters of urban as premise, based on the domestic and foreign theories of urban drainage system, to put forward the construction of storage detention and drainage combined modern urban drainage system. This article takes Shenyang city for example, through the dynamitic programming, applying the theory of dynamitic optimization, planning the optimization urban drainage system of Shenyang city. It can be known after calculation and analysis that the urban drainage system combining with storage detention and drainage has great significance on economics society and ecology. [Nature and Science. 2005;3(2):36-42].

Keywords: urban drainage; rainwater utilization; dynamic optimization theory; Shenyang

1. Introduction

Urban is the central part of politics economics culture and education of an area, also it is the impacted part and the vicissitude of the urban will affect the development of the area.

While about 90 percent of cities in China are by the rivers or seas, all the cities are liable to have flood and water logging disasters. Flood will submerge the cities and destruct the productions and affect the lives. It will take great losses to our country. Because of the special character of flood, almost all the countries have flood problems. With the development of social economic and product forces, more and more people and treasure will be assembled in cities, so if there is a flood and water logging disaster, the losses in cities will greatly surpass those in other areas.

In order to improve the standard of flood control, China has done much control work. All the length of urban levees in China has been 6500km by 1985, which protected cities effectively. While following the construction of flood control engineering and improvement of dispatching capacity, the chances that having water logging disasters are improved owing to the lift of the water level in river course. Urban can't stand up to the impact of water logging disasters.

As we all know, cities of China are short of water. Its per capita water resources are lower than the national standard that is 1000m^3. Over excavating ground water leads to the decline of ground water level and subsidence of ground and formation of the underground funnel. So cities of China have serous flood and water logging disasters and at the same time are short of water resources, adding the importance of protecting environment, it is high time that we put forward a modern urban drainage system.

Modern urban drainage system will combine the flood and water logging disasters elimination with rainwater percolation and the utilization of rainwater. It will decrease the cost of urban drainage system and protect the water quality and environment, and have great significance to the construction of our cities.

2. Theory Of Urban Drainage and Its Calculating Methods

Urban drainage is the drainage of urban rainfall and protection of people's productions and lives from being submerged. It can use lakes, depressions, river channels and pits to detent waterlog, reducing the pressure of urban drainage.

According to the present states and characters of cities in China, it is reasonable to construct the modern urban drainage system. The modern urban drainage system is taking urban dewatering as a part of urban drainage, combining with the constructions of urban depressions, urban rivers etc, formation the urban drainage system that the combination of storage detention and drainage. It will increase water storage space as well as deduce the loss of water logging.

2.1 Urban Dewatering

Urban dewatering is the constructions of drains network of any urban departments, draining the rainstorm water of urban small areas. The task of urban dewatering is collecting and draining the rainfall, preventing from immerging of housing areas and industry and enterprise areas, protecting urban people's lives and treasures. It is rainstorm intensity formula and rainwater discharge formula to derive the designed

rainwater discharge.

2.1.1 Calculated formula of discharge

Designed rainwater discharge is not only the important reference of deciding the profile dimension but also the standard of checking the capability of urban rainwater drains network. Because the area of collecting urban rainwater is small, it often uses the inference formula to calculate the designed rainwater discharge:

$$Q=\varphi F q \qquad (1)$$

There: Q－design rainwater discharge(L/s)

φ－runoff coefficient, it is smaller than one;

F－collecting area（ha）

q－design rainstorm intensity(L/(s • ha))

2.1.2 The formula of rainstorm intensity

Rainstorm intensity formula is the relations among rainstorm intensity I (or q)—rainfall duration t—reoccurrence P. In china the rainfall intensity formula in used is:

$$q=\frac{167A_1（1+c\lg P）}{（t+b）^n} \qquad (2)$$

where: q—design rainfall intensity （ L/(s·ha) ）; p—designed reoccurrence（a）

t—rainfall duration（min）; A_1、c、b、n—local parameter, decided in statistical methods.

2.1.3 Runoff coefficient

Runoff coefficient φ relates to collecting area, slope of ground, density and distribution of constructions, geomorphic state and paving material etc. It is difficult to determine the runoff coefficient accurately. Usually it is empirical value determined by ground covers shown in Table 1.

Table 1. Runoff coefficient

Variety of ground covers	φ	Variety of ground covers	φ
All kinds of roofs and concretes and asphalt pavements	0.90	Dry brick and macadam pavement	0.40
Block stone pavement and macadam pavement treated with asphalt surface	0.60	Not paved earth road surface	0.30
Graded macadam pavement	0.45	Park and grassland	0.15

2.2 Urban rainwater utilization

Urban rainwater utilization is the use of construction and non-construction measures, impounding and utilizing the urban rainwater, reducing the chances of water logging disasters and adding the water resources of urban, improving the water ecological environment, constructing a water loved and water saved society.

There are many kinds of forms that urban rainwater is used, the main forms are infiltration and median water utilization.

(1) Infiltration to supply the ground water

An important method to use rainwater is enforcing the infiltration with all kinds of artificial measures, making more rainwater penetrate into the ground. It can not only improve the urban ecological environment but also reduce the discharge of downstream and decrease the harm of flood and water logging .There are many measure in infiltration. One of them is using permeable material to construct parking and street and pedestrian way ground surface. The second is constructing storage infiltration system: collecting the runoff on the road, after the treatments of device between water and oil and filter, supplying the ground water or irrigating the plants by the roads. It also can use percolation pits and leaky pipes to adding the infiltration.

(2) Median water utilization

Median water utilization is mainly using the roof rainwater which is easy to be collected and has high quality, after treatment it can be used for the supply resources of greening spraying streets etc. Through analyzing the quantity and quality of rainwater, the utilized rainwater quantity can be calculated by the formula:

$$Q=\Psi\cdot\alpha\cdot\beta\cdot A\cdot（H\cdot 10^{-3}） \qquad (3)$$

where: Q—annual average utilized quantity of roof rainwater, m^3;

Ψ—runoff coefficient, it is 0.9;

α—seasonal reduction coefficient;

β—primary waste coefficient;

A—roof level projection area,m^2;

H—annual average rainfall, mm.

Seasonal reduction coefficient α should be divided by analysis of multi-years statistical data from local meteorological department; primary waste coefficient β should be determined by rainfall data and water quality.

3. Research Method of Optimization Plan of Urban Drainage System

There are many measures in improving the urban drainage capacity. Under the condition that attains the designed standard of drainage, how to place the reasonable ratio of all kinds of drainage measures so that when the capacity is given the cost to the measures is the smallest or when the cost is given the capacity to the drainage is the biggest is the problem of optimization plan. Because the relation between engineering measure and cost or engineering measure and drainage capacity is complex, and it is difficult to description. So it has special significance using DDDP to optimize the drainage system.

Although the model of urban drainage system is a one-dimensional problem of DP, its state variable and decide variable is continuous, and its feasible fields are big. If using DP method to study the model, the storage and calculation quantity will surpass the capacity of the computer. So DDDP is the best method to study the model. DDDP method doesn't optimize in all feasible fields but in the some scope of test trace, after successive optimizing, it will attain the best-optimized trace. In the optimization of urban drainage system the DDDP method is often used to simply the calculation.

According to the relation between state variable and decide variable of each phase, we can construct the system formula:

$$S_n = S_{n-1} + x_n(d_n) \quad n = 1,2,3,\ldots N \qquad (4)$$

boundary conditions are $S_0 = 0, S_N \geq W$,

where: W—designed runoff of drainage district.

Adopting sequence succession, the recurrence formula is:

$$\begin{cases} f_n^*(S_n) = \min\limits_{d_n} \left\{ L(S_n, d_n) + f_{n-1}^*(S_{n-1}) \right\} & n = 1,2,3,\ldots N \\ f_0^*(S_0) = 0 \end{cases} \qquad (5)$$

Where: $f_n^*(S_n)$ ——the lowest investment of 1～n stages.

Under the restrict conditions of system equation boundary conditions water balance engineering dimension cost labor and equipments etc, urban drainage system can be optimized. According to the successive equations, urban drainage system can be optimized by phases and we can achieve the best-optimized strategy including engineering dimensions and locations. Because the problem of urban drainage is not a protruding programming, it is needed to seek the best-optimized strategy from different original strategies.

4. Optimization Plan of Urban Drainage System of Shenyang City

4.1 General situation of Shenyang city

Shenyang is the central part of politics economics and culture technology of Liaoning province; and it is the biggest city transportation hinge and material distributing center of northeast region of China. Entire area of Shenyang city is 12,980km^2. The physiognomy of Shenyang gives priority to plain. Shenyang is the transition belt of hill of east of Liaoning province to the alluvial plain of Liao. There are muddy river, thin river, pu river etc. natural rivers and new opening river, such as south canal guard project river, among which muddy river is the main river.

It begins in 1903 that urban drainage network of Shenyang constructed. After many years continue and improvement, there are three rainwater drainage systems, and the all drainage area is 186.97km^2. By the year of 1993, the length of drainage network is 1691km; the drainage network density is 9.14km/km^2, the rate of popularization of it is 88.4%. There are 31 pumping stations in Shenyang including corrupting raining and cloverleaf junction pumping stations. The drainage system of Shenyang is confluence's system or cut-flow confluence's system, and parts of them are distributary's system.

Among the three rainwater drainage systems, there are peace district, Shenhe district, south of Dadong ,part of Dongling and east of Tiexi district Xinghua Street in south drainage system, the all area of it is 74.09km^2; there are west of Tiexi district Xinhua Street and building exploitation section of Shenyang in west drainage system, the all area of it is 44km^2 not including exploitation section of Shenyang; there are Huanggu district north of Dadong district in north drainage system, the all area of it is 68.88 km^2. There are 31 pumping stations in Shenyang drainage system such as Wuai, Gongnong, Congshandong, Congshanxi, Zhonggong, Lianhelu, Zhulin, Jingqin and so on.

The construction of urban drainage system of Shenyang lags to the development of urbanization. During the planning and constructing of city, it takes much important to the new constructions of urban drainage while do less on the inhere urban drainage measures which leads to the non-corollary of the drainage measures and restrict the sustainable development of Shenyang city, and it can not meet the requirement of the market economy development.

4.2 The plan of urban drainage system of Shenyang city

Based on the problems in urban drainage system of Shenyang city and plan thought of urban drainage, we plan the urban drainage system of Shenyang city. On account of there are three urban drainage districts in Shenyang city and the optimization method to each district is the same, we take the south drainage district for example to plan the optimized system of urban drainage. The plan of urban drainage of Shenyang city is shown in Figure 1.

4.2.1 Construction of Mathematics Models of Urban Drainage System

Studying out the urban drainage standard of Shenyang city is half an hour's rainfall to be drained in one hour while the working hour of pumping station is one hour. The reservoir quantum of urban drainage measures is decided by the style of drainage engineering, the material function relation as follows:

Roof rainwater reservoir quantum is $x1 = \Psi q \beta A t d_1$; grassland reservoir quantum is: $x2 = H_2 A d_2$;

Lakes reservoir quantum is $x3 = H_3 A d_3$; river network reservoir quantum is $x4 = H_4 A d_4$;

Pumping station takes out and tonnages quantum is $x5 = 0.36 T_p d_5$;

Where: d_1—the ratio of roof area, d_1 = roof area/all the drainage area，takes as decimal fraction；

d_2—the ratio of grasslands, d_2 = grasslands area/all the drainage area, takes as decimal fraction；

d_3—the ratio of lakes area, d_3 = lakes water surface area/all the drainage area, takes as decimal fraction;

d_4—the ratio of river network area, d_4 = river network water surface area/all the drainage area, takes as decimal fraction;

d_5—designed discharge of pumping station, m^3/s；

A—all the drainage area, km^2;

H_2—storage deep of grasslands, m;

H_3—storage deep of lakes, m;

H_4—storage deep of river networks, m;

T_p—working hour of pumping station in all the drainage period, h;

Ψ—runoff coefficient, there it is 0.9; q—rainfall intensity, L/s·ha;

β—rainfall cut-flow coefficient; there it is 0.8;

t—last period of rainfall, h.

Sequence and reservoir quantum of planed drainage measures in south drainage district of Shenyang city is shown in Table 2.

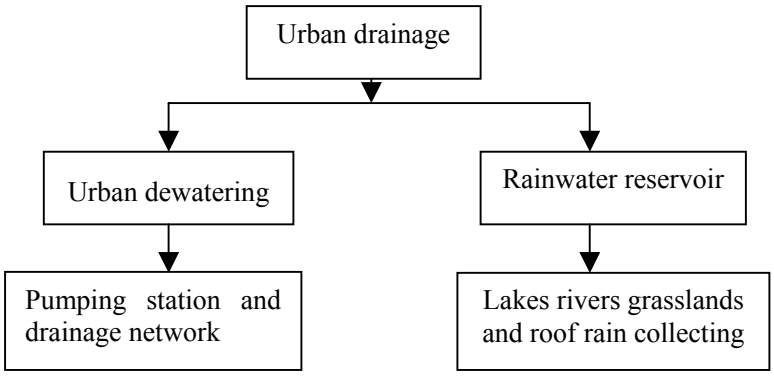


Figure 1. Plan picture of urban drainage of Shenyang city

Table 2. Drainage measurements and engineering investment

Establishment sequence	1	2	3	4	5
Drainage establishment	Roof surface collecting	Grasslands	Lakes	River networks	泵站
Reservoir quantum	$\Psi \cdot q \cdot \beta \cdot A \cdot t \cdot d_1$	$0.3A \cdot d_2$	$0.5A \cdot d_3$	$0.8A \cdot d_4$	$3600 d_5$
Engineering investment (Units)	185 (Yuan/m^2)	30.2 (Yuan/m^2)	90 (Yuan/m^2)	90 (Yuan/m^2)	7 (10^4 Yuan /m^3/s)

System equation is $S_n = S_{n-1} + x_n(d_n)$

n＝1，2，3，4，5

Objection function is $\min F = \sum_{n=1}^{N} L(S_n, d_n)$

Namely min F＝1370665d_1＋223751.8d_2＋666810d_3＋666810d_4＋7d_5 (ten thousand Yuan)

Total runoff Q=φA q, where φ＝0.6，while

$$q = \frac{1825\,(1+0.774\lg P)}{(t_1+2t_2+8)^{0.724}} = \frac{1825\,(1+0.774\lg 0.7)}{(10+2\times 20+8)^{0.724}}$$

=84.93L/s·ha

Therefore Q is 37754.78 m^3/s, while W is 6796×10^4 m^3. Water equilibrium equation is $81.55d_1+2222.7d_2+3704.5d_3+5927.2d_4+0.36d_5=6796$

4.2.2 The Calculation Method of DDDP

Given the allowed precision requirement of reducing increment ε_1＝2% and successive convergent ε_2＝10% and increment of the k time lower than the tenth of elementary increment, we calculate the model as follow steps:

（1）supposing the beginning decision of d（n）, we can calculate the engineering investment of urban drainage system which is shown in Table 3.

（2）Given the beginning increment of {□s（n）}, all the other states are discrete into 3 values except the beginning and terminal states which are fixed value. Supposing the 3 increments of □s_j（n）are 50, 0, -50, then the first iterative results of state fields derived from beginning tracks of s（n）are shown in Table 4.

（3） The first iteration. According to the state points, we can use the DP method to optimize the model. When n=5 n=4 and n=3 the optimized results are shown in Table 5, Table 6 and Table 7. While n=3 and n=2 the method to calculation is the same, so the calculation process is omitted. From Table 8 we can conclude that the minimum engineering investment is F^1=216194.69×10^4 Yuan. Figuring from the given beginning states of s（0）, we can achieve the best-optimized tracks strategies and engineering investment which were shown in Table 9.

Table 3. Drainage engineering investment

Title \ Phases	0	1	2	3	4	5
Beginning decision	0	0.2	0.2	0.08	0.02	16445.2
Beginning states	0	16.31	460.85	757.21	875.754	6796
Engineering investment	500680.76					

Table 4. The state points of the first iteration process

States \ Phases	0	1	2	3	4	5
s（n）＋□s_1（n）	0	66.31	510.85	807.21	925.754	6796
s（n）＋□s_2（n）	0	16.31	460.85	757.21	875.754	6796
s（n）＋□s_3（n）	0	0	410.85	707.21	825.754	6796

Table 5. Values of optimization on n=5

Sequence of netting points	S(4)	S(5)	d(5)	F(5)	$d^{※}$(5)
S(4)+△S_1(n)	925.754	6796	16306.24	114143.67	
S(4)+△S_2(n)	875.754	6796	16445.13	115115.89	
S(4)+△S_3(n)	825.754	6796	16584.02	116088.12	

notes: $d(5) = [s(5) - s(4)] \times 10^4 / 3600; f(5) = 7 \times d_5 \times 10^4 (Yuan)$

Table 6. Values of optimization on n=4

Sequence of netting points	S(0)	S(1)	D(1)	f(1)	$\sum_{i=1\sim5} f(i)$	$d^{※}(1)$	$d^{※}(2)$	$d^{※}(3)$	$d^{※}(4)$	$d^{※}(5)$
0	0	66.31	0.813	1114350.6	1323870.13					
0	0	16.31	0.20	274131.89	488684.702					
0	0	0	0	0	216194.69	0	0.23	0.053	0.02	16584.02

Notes: $d(1)=[s(1)-s(0)]\times10^5/(0.9\times74.09\times84.93\times1800\times0.8)$ $f(1)=185\times74.09\times10^2 d_1\times10^4(Yuan)$

note: $d(4)=[s(4)-s(3)]/(0.8\times74.09\times10^2)$ $f(4)=90\times74.09\times10^2 d_4\times10^4(Yuan)$

Table 7. Values of optimization on n=1

Sequence of netting points	S(3)	S(4)	d(4)	f(4)	f(4)+f(5)	$d^{※}(4)$	$d^{※}(5)$
S(3)+△S_1(n)	807.21	925.754	0.02	13336.2	127479.87		
S(3)+△S_1(n)	807.21	875.754	0.012	8001.72	123117.61		
S(3)+△S_1(n)	807.21	825.754	0.003	200.43	118088.55	0.003	16584.02
S(3)+△S_2(n)	757.21	925.754	0.028	18961.2	133104.87		
S(3)+△S_2(n)	757.21	875.754	0.02	13336.2	128452.09		
S(3)+△S_2(n)	757.21	825.754	0.012	7711.2	123799.32	0.012	16584.02
S(3)+△S_3(n)	707.21	925.754	0.037	24586.2	138729.87		
S(3)+△S_3(n)	707.21	875.754	0.028	18961.2	134077.09		
S(3)+△S_3(n)	707.21	825.754	0.02	13336.2	129424.32	0.02	16584.02

Table 8. The optimized states strategies and engineering investment

Titles \ Phases	0	1	2	3	4	5
s(n)	0	0	510.85	707.21	825.754	6796
d(n)	0	0	0.23	0.053	0.02	16584.02
Engineering investment	216194.69					

Table 9. The optimized strategies

Titles \ Phases	0	1	2	3	4	5
Optimized decision	0	0	0.2467	0.0328	0.02	16688
Optimized state	0	0	548.4	669.7	788.3	6796
Engineering investment	207200					

Table 10. Optimized values on different beginning

Sequence of beginning strategies	d1	d2	d3	d4	d5	$d^{※}1$	$D^{※}2$	$d^{※}3$	$d^{※}4$	$d^{※}5$	F_{min}
1	0.2	0.2	0.08	0.02	16445.2	0	0.2467	0.0328	0.02	16688	207200
2	0.1	0.1	0.05	0.04	17064.5	0	0.1430	0.0028	0.04	17308	181670
3	0.4	0.3	0.1	0.1	14259.4	0	0.3	0.0818	0.1	14502	291130

(4) Take the first improved iteration trace $\{s^1(n)\}$ as the second iteration test trace. Comparing the objection function values attained before and after iterations, we can know if it meets the condition to

reduce the increment.

$$\frac{|F^1-F^0|}{F^0}=\frac{|216194.69-500680.76|}{500680.76}=56.81\%>2\%,$$

So we do not reduce the increment.

(5) The second iteration. Firstly rounding $\{s^2(n)\}$ we construct new state points by the original increment with the same method as the first iteration. Then we do the second optimized iteration with DP method on the new state points. Though iterating we can get new improved trace $\{s^2(n)\}$ new improved strategy $\{d^2(n)\}$ and new corresponding optimized objection function value F^2. Comparing the objection function values before and after iterations, we can conclude that if it doesn't meet the condition that reducing increment we will continue iteration at the original increment, while if it meets the condition, we will construct new corridor with the descending increment, which is half of the original increment and continue iterating calculation. When it meets the condition of reducing increment again we will reduce increment again. At the k time when iterating increment meets the condition that $\{\Delta s^k(n)\}\leq 0.1\{\Delta s^1(n)\}$ the iteration calculation ends.

Through the discrete differential dynamic planning program in Matlab, after three iterations we can attain the best-optimized trace and best-optimized strategy that are shown in Table 9.

(6) Because this problem not always to be protruding plan, we should optimize the strategy from different beginning strategies. The calculating results are shown in Table 10.

4.2.3 Results

After optimized calculation we can conclude that under the condition of obtaining the standard of urban drainage when d_1 gets 0, d_2 gets 0.1430, d_3 gets 0.0028, d_4 gets 0.04, d_5 gets 17308 the urban drainage system gets the lowest investment that is 181670×10^4 Yuan.

5. Conclusions

This plan aims at reducing the pressure of urban drainage, combining flood control and draining water logging with the utilization of rainwater resources. It is in favor of the sustainable utilization and management of water resources and the sustainable development of social economics.

When the plan brought into effect, it will decrease the runoff effectively and delay the time of conflux and increase the capacity of drainage protecting people's lives and property adding the supplement for groundwater and in some degree relaxing the water resource crisis of Shenyang city. It can also lessen the surpass water on the roads and reduce the spread of contamination with water, improving city environment.

The combination of utilization of rainwater and engineering of gardens makes Shenyang city watery and green and creates a graceful living conditions for Shenyang city people.

Correspondence to:
Junshi He
College of Water Resources
Shenyang Agricultural University
Shenyang, Liaoning 110161, China
Telephone: 01186-13998386658
01186-24-88487134
E-mail: hejunshi@163.com

References

[1] Chen bin, Study on Urban drainage system. Water and Electricity of Chinese Country. 1997;2:33-5.
[2] Chen Yuheng, Conceive the using of urban flood. Water Conservation Development Research. 2002;4:32-3.
[3] Liu Shukun, Pilot study on water conservation modernization of China. Water conservation development research.2002, 2(12):7-11.
[4] Yu Qiyang, Enhance the using of urban flood resources in other to ensure urban sustainable development. Water conservation development research.2002, 2(3):12-3
[5] Sun Huixiu, Drainage project. Chinese Building Industry Publishing Company. Beijing. 1999.

Studies on the Optimal Process to Extract Flavonoids from Red-raspberry Fruits

Yaqin Xu, Rui Zhang, Hong Fu

Science College, Northeast Agriculture University, Harbin, Heilongjiang 150030, China, Xu-yaqin@163.com

Abstract: Effects of single factors such as temperature, time, material ratio and concentration of ethanol solution on the contents of flavonoids were investigated. On this basis, employing the orthogonal design, the optimum extraction conditions for flavonoids from raspberry fruits were determined by colorimetric estimation. Experiment results: the amount of flavonoids extract reached its maxima when extracted for 3 h by using 95% ethanol solution at 80°C with the ratio of raw materials to ethanol solution of 1:10. [Nature and Science. 2005;3(2):43-46].

Key words: red-raspberry; flavonoids; extract; single-factor experiments; orthogonal experiments

1 Introduction

Red-raspberry, which belongs to genus Rubus L. of family to *Rasaceae*, is a kind of berry with tender and juicy fruits. Fruits of red-raspberry are rich in vitamins and minerals and especially abundant in caner- fighting, anti- aging substances such as SOD and gallogen (Häkkinen, 2000; Puupponen- Pimiä, 2005), therefore famous as "Fruit of Life" in the US. Apart from those, raspberry fruits contain significant flavonoids. It was reported that flavonoids could remove the O_2^- • in human bodies, strengthen the natural disease-fighting system, improve blood circulation and lower blood pressure (Fang, 1998; Liu, 2002; Wang, 1996). At present, studies on the extraction of flavonoids from raspberry fruits have not been reported. In this study the Optimum conditions to extract Flavonoids from Red-raspberry fruits were studied systematically in order to achieve scientific evidence for the processing and utilizing of raspberry fruits.

2 Materials and Methods

2.1 Materials

The raspberry fruits were collected from the Horticultural Center of Northeast Agricultural University and put into air-tight containers to store in a refrigerator after instantly frozen and minced.

2.2 Extracting process

At first the main factors of the extraction temperature, the concentration of the extracting agent, extraction time and the materials ratio (the weight of fruits: volume of the extracting agent), which affect the extraction of flavonoids were studied individually, and then the optimum extracting conditions of flavonoids from raspberry fruits were determined by adopting $L9(3^3)$ orthogonal experiments after making analysis of variance and significance evaluation diagram of extracting process was: materials→extract →filter →enrichment → extracted by using ethyl acetate→raw extract of flavonoids→confirm the volume by using 95% ethanol solution→ measure.

2.3 Measuring of the contents of flavonoids

The contents of flavonoids was measured using colour comparing method of $NaNO_2$-$Al(NO_3)_3$ and calculated using the following formula (Xu, 2000):

Relative contents of flavonoids (μ g/g)=5YV/W

Y= 75.944A- 0.3494

A: absorbance

W: the weight of raspberry fruits precisely measured:g

V: the volume of extracting agent:mL.

3 Results and Discussions

3.1 Effects of single factor for flavonoids extraction

3.1.1 The effect of the concentration of ethanol solution on the content of flavonoids extracted

The result of Figure 1 showed that the content of raw flavonoids extract increases with the concentration of ethanol solution and the contents increased slowly when the concent- ration was greater than 75%. However, raw flavonoids extract had better solubility and the solvent was easily vaporized and reused in 95% ethanol solution, 95% ethanol solution was used.

3.1.2 The effect of temperature on the contents of flavonoids extracted

Figure 2 was contents of raw flavonoids extracted at 50°C, 65°C, 80°C, 95°C under certain conditions.

The contents of raw flavonoids extract tended to increase gradually with the rise of temperature in range of 50°C~80°C. It may be probable that the greater speed

of the molecule movements in higher temperature so that flavonoids diffused more quickly from cell to extracting agent. But the flavonoids could be oxidized at temperature of surpassing 80° so that the contents of flavonoids extracted started to decrease gradually. 80°C was the optimum temperature.

3.1.3 The effects of the material ratio on the contents of flavonoids extracted

The result of Figure 3 showed that the contents of flavonoids increased significantly with the rise of the material ratio in a range of 1:5~1:10 and changed unobviously when the ratio was greater than 1:10. This was because when the material ratio was between 1:5 and 1:10, flavonoids were extracted fully with the rise of volume the extracting agent so as to the contents of flavonoids increased. Otherwise, when the material ratio reached a certain level, the extract has well soluted in the solution lead to the contents of extract become steady and wouldn't increase significantly. The optimum material ratio was 1:10.

3.1.4 Effect of flavonoid extraction time

The result of Figure 4 showed contents of flavonoids extracted for 3 h reached its maxima and then decreased with prolonging extraction time. This may increase amount of impurities with the rise of time. Therefore the optimum extraction time was 3 h.

3.2 Optimum conditions for flavonoid extraction

On the basis of the effects of single factor in above chapters, adopting extracting agent of 95% ethanol solution, the trials of tri- factor, tri-level Orthogonal design were conducted and the results were given in Table 1 and Table 2.

The results of orthogonal experiment were made an analysis of range, single-factor analysis of variance and the analysis of linear regression by SPSS software. Table 2, Table 3 and Table 4 showed the results.

The results of all kinds of analysis showed the degree of affecting the contents of flavonoid extract was C>B>A. The material ratio had the greatest effect on the test results, time of extraction rank second and the temperature came last. When extraction temperature and extraction time were at the 2nd level, the contents of extract reached its maxima. When the material ratio was at the 3rd level, the amount of extract reached its maxima. However, to save the solvent, it's not economical to adopt for the material ratio of 1:15 even it got the greatest contents of extract. Table 2 showed that the 2nd level was suitable for all three factors because of the greater contents of extract than the 1st level.

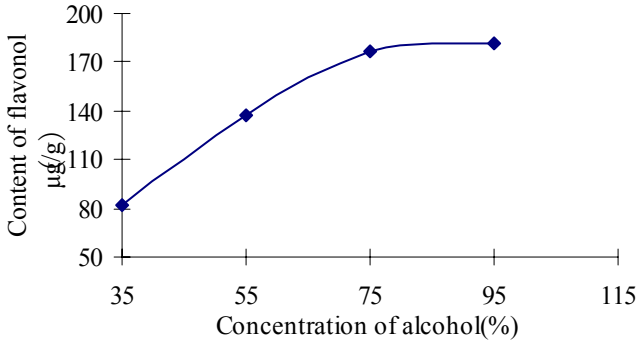


Fig.1 Effect of different concentration of alcohol

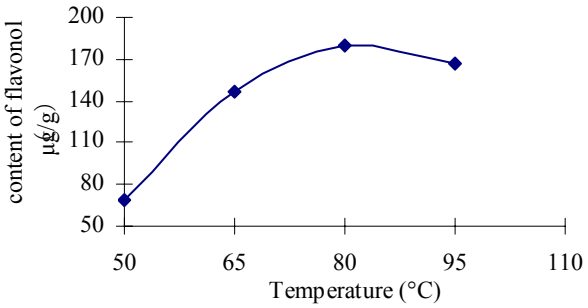


Figure 2. Effect of different temperature

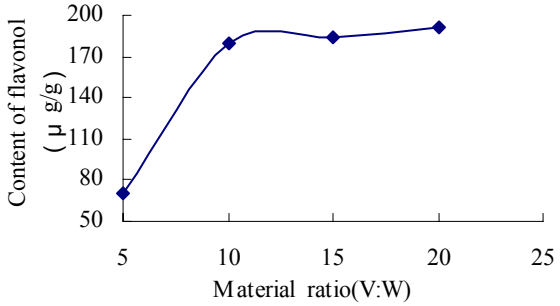


Fig.3 Effect of the material ratio

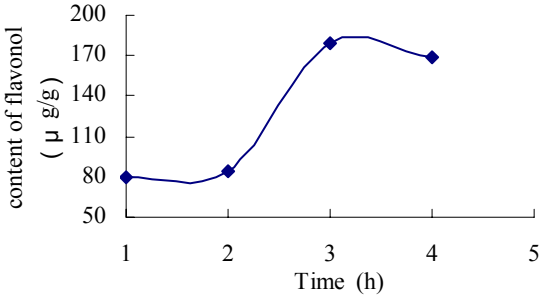


Fig.4 Effect of the different extraction time

Table 1. Factors and levels

Level	Factor		
	A. Temperature（℃）	B. Extraction time (min)	C. Ratio of solid to liquid (W:V)
1	65	120	1： 5
2	80	180	1：10
3	95	240	1：15

Table 2. Experiment results and range analysis

Item	A	B	C	Flavonoids (μ g/g)
1	1	1	1	80.2
2	1	2	2	157.4
3	1	3	3	164.2
4	2	1	3	143.3
5	2	2	1	107.2
6	2	3	2	179.8
7	3	1	2	142.2
8	3	2	3	182.8
9	3	3	1	100.0
K_1	401.8	365.7	287.5	
K_2	430.3	447.4	479.3	
K_3	425.0	444.0	490.3	
k_1	133.9	121.9	95.8	
k_2	143.4	149.1	159.8	
k_3	141.7	148.0	163.4	
R	9.5	27.2	67.6	

Table 3. Single- factor analysis of variance

	Sum of Squares	df	Mean Square	F	P
A	153.3	2	76.6	0.044	$F_{0.01}$（2.2）=99
B	1423.0	2	711.5	0.467	$F_{0.05}$（2.2）=19
C	8670.9	2	4335.4	13.788	$F_{0.1}$（2.2）=9
Total	10247.2	6			$F_{0.25}$（2.2）=3

Table 4. Analysis of variance in linear regression

Model		Sum of Squares	df	Mean Square	F
1	Regression	7965.613	3	2655.204	5.122
	Residual	2591.897	5	518.379	
	Total	10557.510	8		

4 Conclusion

To sum up, the Optimum process to extract Flavonoids from raspberry fruits were obtained, namely Extracted for 3 h by using 95% ethanol solution at 80°C with the material ratio of 1:10 (W:V).

Correspondence to:
Yaqin Xu
59 Wood Street
Gongbin Road, Xifang District
Harbin, Heilongjiang 150030, China
Telephone: 01186-451-5519-0732
Cellular phone: 01186-130-7453-5378
E-mail: xu-yaqin@163.com

References

[1] Fang P. Biological Effects of Flavones. Journal of Dali medical college 1998;7(4):52~54.
[2] Häkkinen SH, Kärenlampi SO, Heinonen IM, Törrönen AR. Ellagic Acid Content in Berries: Influence of Domestic Processing and Storage. European Food Research and Technology 2000;212(1):75~80.
[3] Liu M, Li X. Antioxidant and Antiprolife -rative Activities of Raspberries. J Agric Food Chem 2002;50:2926~2930.
[4] Puupponen-Pimiä R, Nohynek L, Alakomi HL, Oksman-Caldentey KM. Bioactive Berry Compounds-novel Tools against Human Pathogens. Applied Micro-biology and Biotechnology. 2005;67(1):8~18.
[5] Xu Y, Sun Y, Fu H. Studies on the Optimal extraction Conditions of Flavonoid in Leaves of Black Currant. Journal of Northeast Agriculture University 2000;7(2):144~146.
[6] Wang H, Cao G, Prior R. Total antioxidant capacity of fruits. J Agric Food Chem 1996;44:701~705.

A 3D Crotched Double-band Dipole Antenna with Wide Impedance Bandwidth Novel Design for WLAN System

Hsien-Chiao Teng,[1], Chih-Sen Hsieh[2], Shen Cherng[3], Yuan-Tung Cheng[2]

[1]Department of Electrical Engineering, Chinese Military Academy, Fengshan, Taiwan 833, ROC

[2]Department of Electronic Engineering, Chengshiu University, Niaosong, Taiwan 833, ROC

[3]Department of Electrical Engineering, Chengshiu University, Niaosong, Taiwan 833, ROC

Abstract: A novel double-band 3D crotched dipole antenna is introduced. This crotched dipole antenna has dual-band operation. A prototype, which is suitable for application in wireless local area network (WLAN) design in 2.4 GHz and 5 GHz bands, is characterized by wide impedance bandwidths and constructed for laptop as an internal antenna. The L-shaped microstrip feed line structure effectively increases the impedance bandwidth to 11% in lower band covering the 2.4 GHz band (2400–2484 MHz) and 50% in upper band covering the 5.2 GHz (5150–5350 MHz) and 5.8 GHz (5725–5875 MHz) bands for the proposed dipole antenna. [Nature and Science. 2005;3(2):47-51].

Key words: dipole antenna; WLAN; impedance bandwidth

1 Introduction

Compact and wide impedance bandwidth design techniques for planar antennas have been very much concerned recently. In addition, researchers are paying attention in planar antennas for wireless local area network (WLAN) systems in the 2.5 GHZ and 5.2 GHz bands. The presented 3D crotched dipole antenna in this letter is desired to put 3D crotched dipole antenna in use on the WLAN and characterized by wide impedance bandwidth to match the limitation of the inner dimension of the laptop. By the result of the experimentation, the radiator patch only needs 5 mm in height to cover 2.4 GHz (2400-2483MHz), 5.2 GHz (5150-5350MHz) and 5.8 GHz (5725-5875MHz) bands for WLAN operation. The geometry of 3D-L shaped two sides microstrip crotched dipole design also enhanced the antenna gains across the lower band and the upper band attractively for practical application.

2 Antenna Design

Figure 1 shows the configuration of the 3D crotched antenna fed by a 1 mm wide and 5 mm in length microstrip feed line on feeding substrate. In Figure 2, it revels the wide impedance bandwidth covering 2.4 GHz and 5 GHz double-band operations for the proposed 3D crotched antenna. The fabrication of the proposed antenna uses the FR4 substrate. All of the system circuit can be fitted in with laptop device. The dimension of the ground plan is 1×7.4 mm² and the perpendicular antenna plan is 5×44 mm². The 50 Ω transmission line was pre-designed 1×7 mm² and simulated by IE3D simulator. The total length of the antenna is 44 mm. The positive arm connected to a 4×10 mm^2 rectangle plane is 23 mm in length and can be excited at 2.4 GHz. The crotched radiator system is 10 mm in length and can be excited at 5 GHz. The width of all line segments is 1 mm. The crotched short circuit connecting to the ground plane established an additional inductance to compensate for the large capacitance arising from the area between the crotched radiator and the ground plane.

3 Experimental Results and Discussion

The measured return loss agrees with the simulated result obtained with the software high-frequency

structure simulator (HFSS) simulator. Figures 3, 4 and 5 are the measured radiation patterns at 2448, 5342 and 5726 MHz, respectively. In the azimuthally plane (x–y plane), the E-component shows small variations, which indicates that stable radiation patterns are obtained in every operating band. Figure 6 shows the measured peak antenna gain. For the 2.4 GHz band, a stable peak antenna gain is measured about 3.0 dBi. For the 5.3 and 5.7 GHz bands, the peak antenna gain ranges from about 5.8–6.2 dBi, which is attractive for practical use in higher band operation with reduced-size patch antenna.

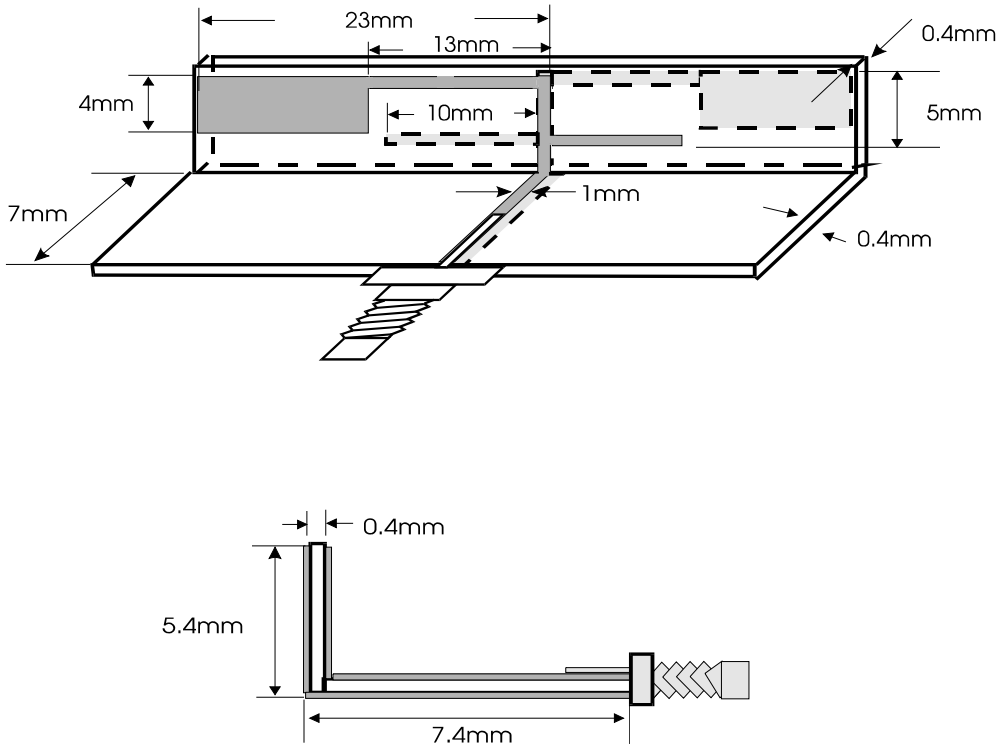


Figure 1. The proposed geometry of double-band 3D crotched dipole antenna.

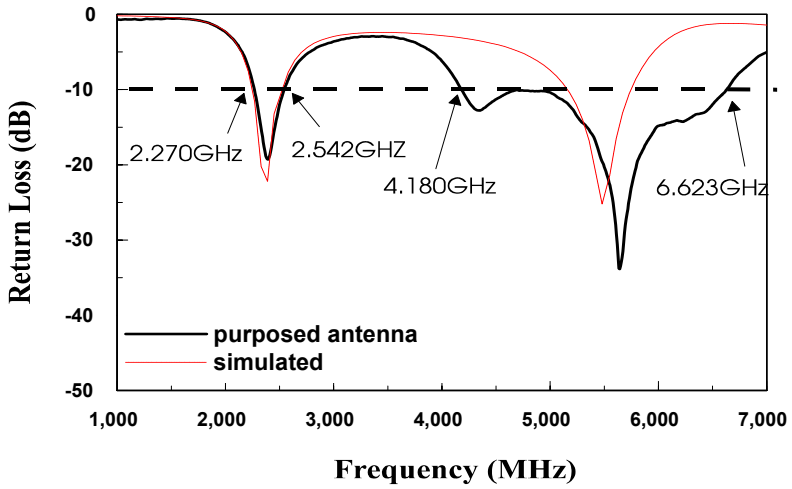


Figure 2. Measured and simulated return loss.

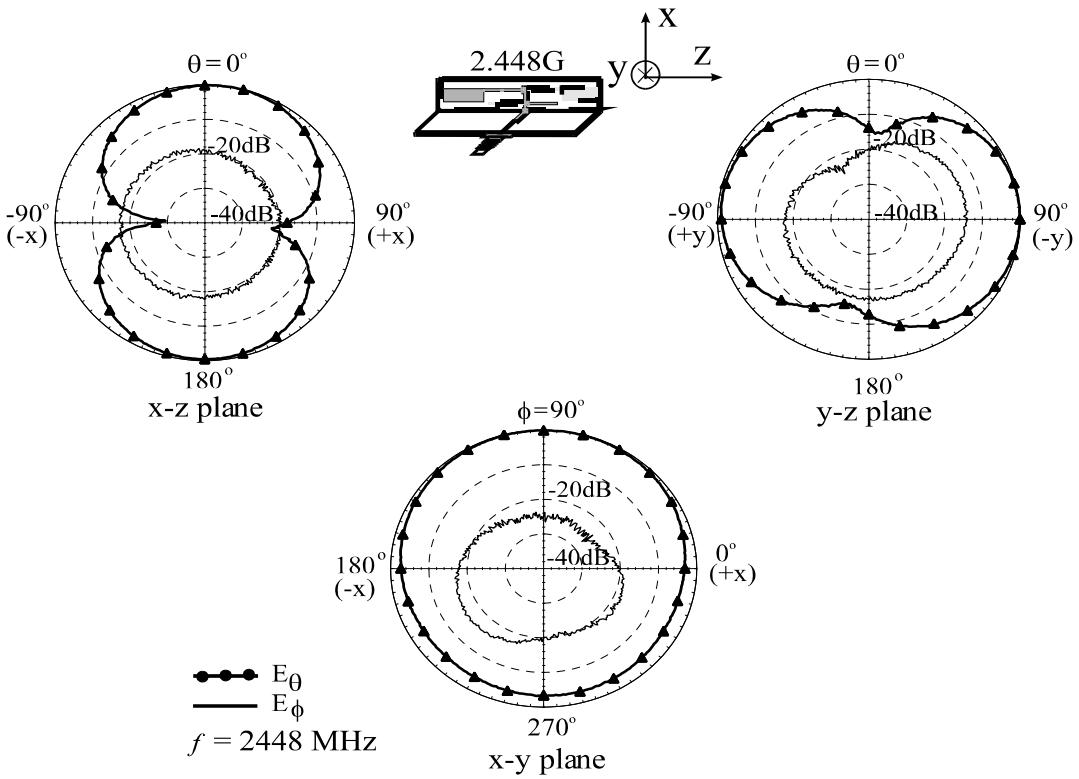


Figure 3. Measured radiation patterns at 2448 MHz.

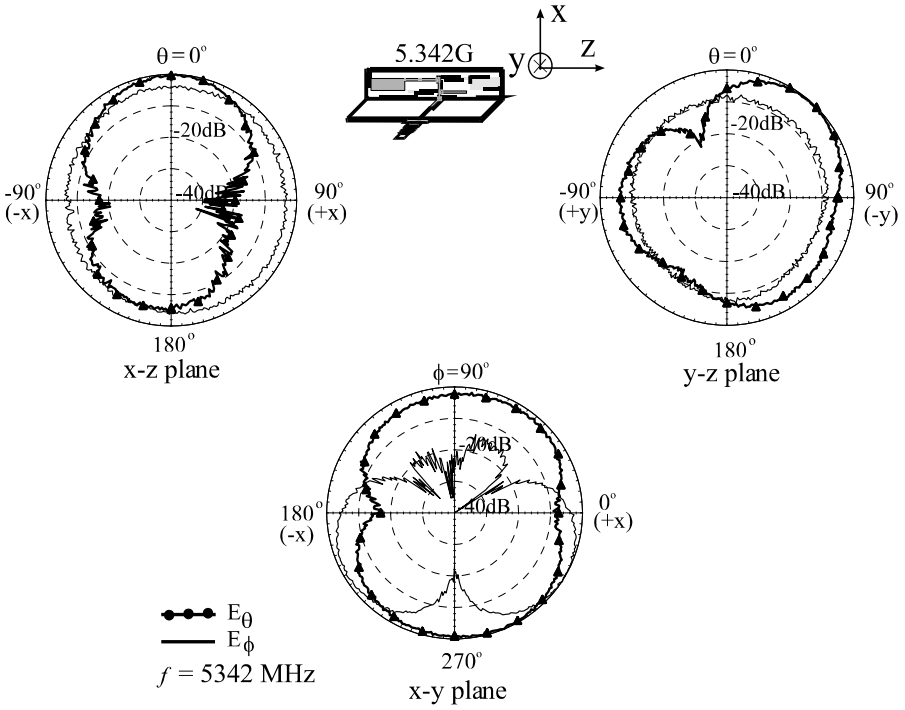


Figure 4. Measured radiation patterns at 5342 MHz.

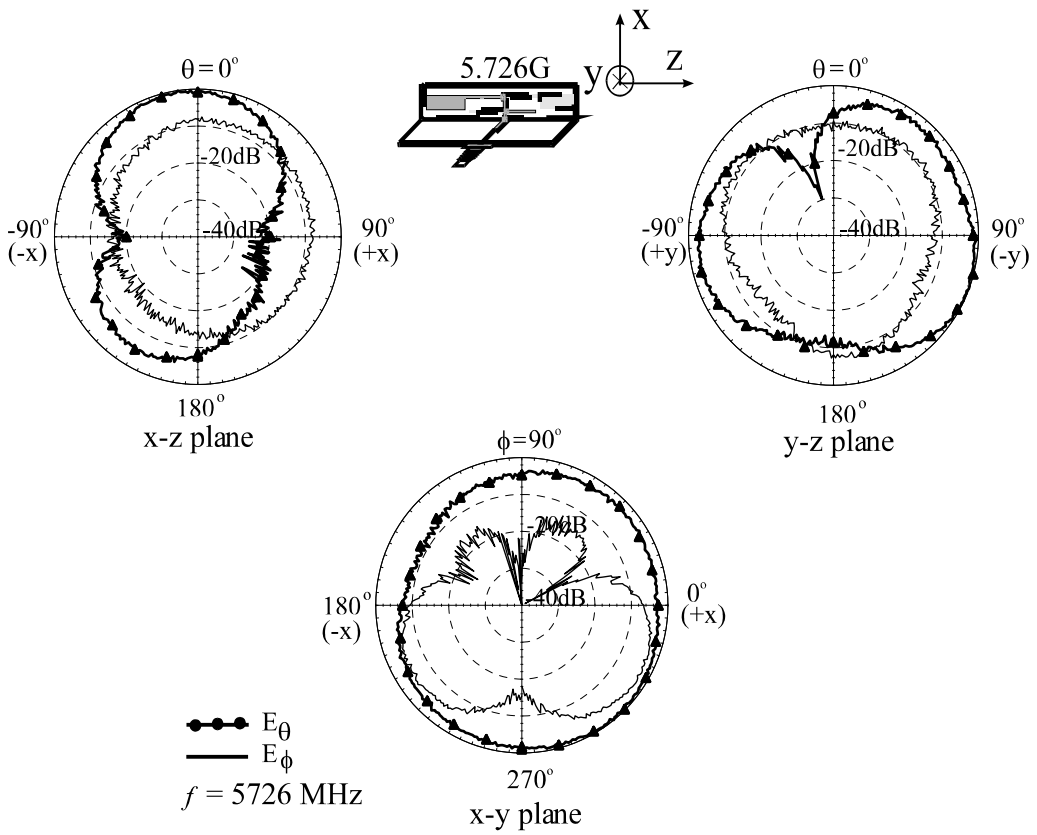


Figure 5. Measured radiation patterns at 5726 MHz.

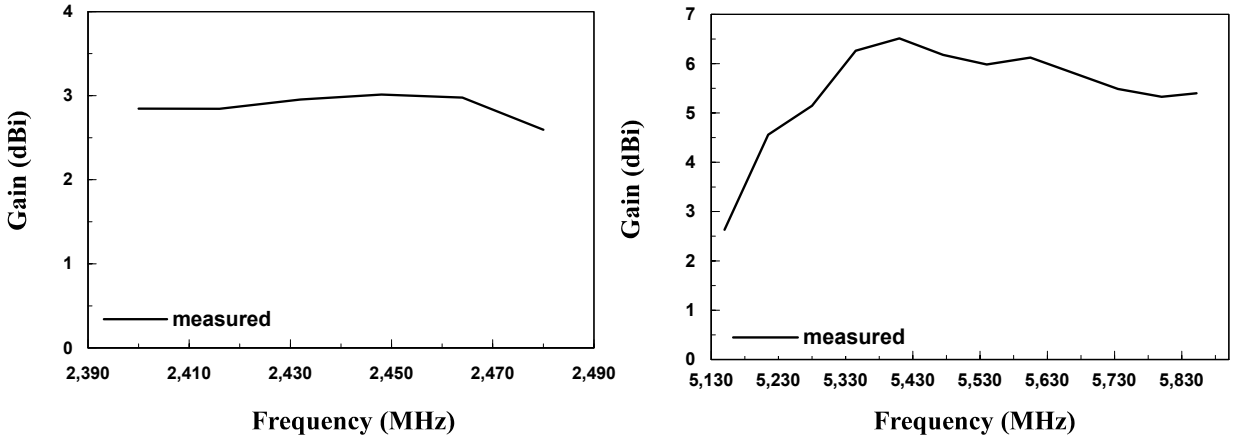


Figure 6. Measured peak antenna gain. (a) The 2.4 GHz band. (b) The 5.3/5.7 GHz bands.

4 Conclusions

The geometry structure of novel 3D crotched antenna is simply L shaped [2] with 5 mm in height, which is suitable for the size of laptop and provides attractive antenna gains across the dual band in operating frequencies at different excited modes. We are successful to design an antenna, which is applicable for WLAN with wide impedance bandwidth, 11 % in lower band and 50 % in higher band, and attractive antenna gains, 3.0 dBi across the lower band and 6.0 dBi across the upper band as well. In addition, L-microstrip line feed structure effectively reduces the difficulty to compromise between the inducing capacitance of feed substrate and inductance of the probe.

Correspondence to:
Hsien-Chiao Teng
Department of Electrical Engineering, Chinese Military Academy, Fengshan, Taiwan 833, Republic Of China
Telephone: 011886-7747-9510 ext 134
Email: scteng@cc.cma.edu.tw

References

[1] Wong KL, Su SW, Kuo YL. A printed ultra-wideband diversity monopole antenna. Microwave Opt Technol Lett 2003;38:257-9.
[2] Xu P, Fujimoto K. L-shaped self -complementary antenna, in 2003 IEEE Antennas Propagat Soc Int Symp Dig 2003;3:22-7.

Application of Genetic Algorithm（GA）Trained Artificial Neural Network to Identify Tomatoes with Physiological Diseases

Junlong Fang，Shuwen Wang, Changli Zhang

Engineering College, Northeast Agricultural University, Harbin, Heilongjiang 150030, China, swanhaha@163.com

Abstract: We synthetically applied computer vision, genetic algorithm and artificial neural network technology to automatically identify the tomatoes that had physiological diseases. Firstly, the tomatoes' images were captured through a computer vision system. Then to identify cavernous tomatoes, we analyzed the roundness and detected deformed tomatoes by applying the variation of fruit's diameter. Secondly, we used a Genetic Algorithm (GA) trained artificial neural network. Experiments show that the above methods can accurately identify tomatoes' shapes and meet requests of classification; the accuracy rate for the identification for tomatoes with physiological diseases was up to 100%. [Nature and Science. 2005;3(2):52-58].

Key words: tomato with physiological disease; computer vision; artificial neural network; genetic algorithms

Introduction

In China, most traditional fruit quality tests still remain at the preliminary stage in which identification is determined by human sense organs. These subjective evaluations are affected by personal abilities, color resolution, emotions, fatigue and rays of light etc. Therefore these techniques are inefficient and inaccurate, and most of them are observational appreciations whose objectivity and accuracy are inadequate. The result is that China's export fruits are of inferior external appearances, and lack competitive power in international markets. Thus it is rather indispensable to improve the national test level of fruit quality[1]. Contrasted with other test technologies, the advantages of computer vision include high speed, multiple functions, capability to do assignments that man or mechanical grading machines are incompetent to do, and the ability to automatically achieve single time identifications of shape, size, degree of maturation, and surface flaws; all of which greatly help increase grading precision[2]. Early in 1985, Sarkar and Wolfe et al successfully developed a kind of quality grading equipment for fresh tomatoes, using computer vision with an applicable lighting fixture and direction finding mechanism. Its grading error is only 3.5%, far superior to the grading precision of manual tests, although still low speed.

1 Capture of the tomato's image

The tomato's image is captured by a computer vision system, which consists of computers, image acquisition cards, CCD (Charge Coupling Device) cameras, 20-watt annular incandescent lamps, etc. Because of the tomato's extremely-smooth surface, illumination results in serious specular reflection. It was found that if a soft-light sheet was added between a 20-watt annular incandescent lamp and the tomato, the tomato's image would no longer contain specular reflection[5]. Figure 1 shows the tomato's acquired image.

2 Network's establishment

2.1 Brief introduction to Genetic Algorithm

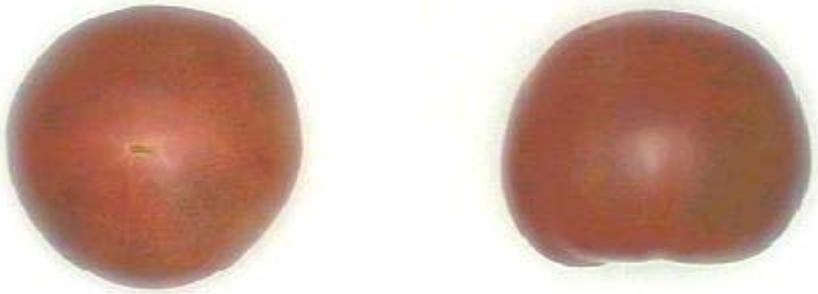
Genetic Algorithm (GA) is a kind of search and optimized algorithm that have been produced from simulating biologic heredities and long evolutionary processes of creatures. It stimulates the mechanism of "survival competitions; the superior survive while the inferior are eliminated, the fittest survive." The mechanism searches after the optimal subject by means of a successive iterative algorithm. Ever since the late 80s, GA, as a new cross discipline which has drawn

people's attention, has already shown its increasing vitality in many fields[6].

GA stimulates reproduction, mating, and dissociation in natural selection and natural heredity procedures. Each possible solution to problems is taken as an individual among population, and each individual is coded as character string; each individual is evaluated in response to predefined objective functions and a flexibility value given. Three of its elemental operators are selection, crossing, and mutagenesis[7].

Its main features are as follows:

(1) GA is to acquire the optimal solution or quasi-optimal ones through a generational search rather than a one-point search; (2) GA is capable of global optimum searching; (3) GA is a parallel process to population change, and provides intrinsic parallelism; (4) The processed object of GA is the individuals whose parameter set is coded rather than the parameters themselves, and this very feature enables GA to be used extensively[8].



(a) 0^0 image (b) 90^0 image

Fig.1 Tomato's image

2.2 Model establishment

1) Determination of the genetic encoding scheme

The first step, when GA is applied, is the determination of a genetic encoding scheme, namely to denote each possible point in the problem's search space as a characteristic string of defined length. This is in order to be sure that GA will not only optimize network configuration but, in the meantime, genetic training will proceed on weight values. In this paper, weight values between each layer of the multi-layer feed-forward neural network are simultaneously coded as one chromosome.

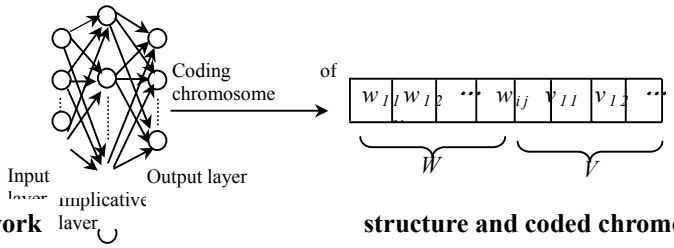


Figure 2. Neuron network structure and coded chromosome

Structure and weight value of the neuron network are mapped to this very chromosome, one of whose layers is the implicative layer whose nodes are determined in accordance with actual demands. Both W as the weight array from the input layer to the implicative layer and V as that from the implicative

layer to the output layer are still arranged in row-to-column sequence and binary-coded. Although such genetic coding helps increase the length of chromosome string, yet each gene value is strictly simplified to either "0" or "1", which remarkably improves the algorithm's global search ability.

2) Definition of flexibility function and its transformation of scale

The second key step, when GA is applied to generate a neural network, is the definition of the flexibility function to evaluate the problem-solving ability of the neural network, which is denoted by a certain specific chromosome string. In this paper, objective function is generated from the cost factor *mean-square error* (*MSE*) of the neural network output, and then converted into a function via reciprocal transformation. Its computational formula is as follows:

$$MSE=\frac{1}{mp}\sum_{P=1}^{p}\sum_{j=1}^{m}(d_{pj}-y_{pj})^2$$

where; m=sum of output nodes; p=sum of trained samples; d_{pj}=expected output of network; y_{pj}=actual output of network.

In GA's later search period, sufficient varieties will probably still exist among population, whereas population's average flexibility value may approximate the optimal variation. If such status isn't changed, then amongst subsequent generations, the individuals that have average value and those optimal ones will approach identical reproduction quantities, so actually here almost no competition exists, which consequently slows down the convergence rapidity.

To avoid the above situation, reproduction quantity in the neural network evolutional process will be adjusted to improve algorithmic performance. Reciprocal scale transformation to the objective function is done as:

$$f(MSE) = \frac{K}{MSE}$$

in which the selection of K is quite crucial, as it determines the coerciveness of the selection procedure. K is determined from experiments.

3) Three of GA's elemental operators

(1) Selection. In the paper, the disk gambling method is adopted, and calculated as

$$P_i = \frac{f_i}{\sum_{i=1}^{N} f_i} = \frac{f_i}{f_{sum}}$$

in which; f_i=flexibility value of individual i; f_{sum}=total flexibility value of population; P_i=selective probability of individual.

It's obvious that individuals with high flexibility values are more likely to be reproduced[9].

(2) Crossing. In the paper, one-point crossing is adopted. The specific operation is to randomly set one crossing point among individual strings. When crossing is executed, partial configurations of the very point's anterior and posterior individuals are exchanged, and give birth to two new individuals.

(3) Mutagenesis. As for two-value code strings, mutagenic operation is to reverse the gene values on the gene bed, namely covert 1 to 0 or 0 to 1.

4) **Genetic control parameters**

Genetic control parameters in the paper are chosen as follows: N as population scale is 50, P_c as crossing probability is 0.4, P_m as mutagenic probability is 0.003, and network's terminative condition is $MSE \leq 0.05$.

3 Extraction of characters

Deformed fruits and cavernous ones result from two major physiological diseases. Cavernous fruit refers to ones with corner angles on their surfaces, and its tangential faces are mostly triangle in shape (Figure 3a). The chief characteristic of deformed fruits is its irregular shapes, such as abnormal, fingerlike projections or papillary (Figure 3b) fruitages, etc. Furthermore, the cavernous fruit can be identified through roundness value of 0° image, and the deformed one identified through variations of the fruit's diameter of 90º image.

(a) cavernous tomato (b) abnormal tomato

Figure 3. Tomatoes with physiological diseases

3.1 Roundness character

Roundness, which is derived from the area and the perimeter, measures the shape complexity. Its formula is[10]

$$e = \frac{4\pi A}{p^2}$$

in which A—the area; P—the perimeter.

It can be seen that, e=1 if round in shape. Since tangential faces of the cavernous fruit are triangular, its roundness is less than 1.

3.2 Variations of fruit's diameter

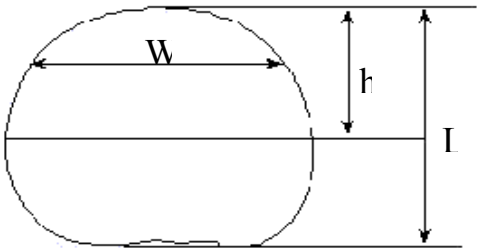


Figure 4. Characteristic parameters of tomato's shape

Variations of the fruit's diameter refers to the variations of diameter from tomato's top to its bottom [11] (Figure 4), which can be defined as the distance W_i between the two intersections derived from a line perpendicular to fruit's axis and tomato's edge. In this way, the tomato is divided perpendicular to fruit's axis into twelve equal parts, then W_1--W_{12} are extracted as characteristic parameters of variations of the fruit's diameter. Due to the unsmooth variations of its diameter, deformed fruits are detected by this means.

4 To identify tomatoes with physiological diseases

4.1 To identify cavernous tomatoes

Roundness is calculated via the perimeter P and the area. P is measured in millimeters, and A is measured in square millimeters, while both the length and the area in the plane coordinate system of image are measured in pixels. The quantitative relationship obtained from experiments of the tomato's perimeter in pixels at maximum transverse diameter is shown in Table 1, and the quantitative relationship of the tomato's area at maximum transverse diameter is shown in Table 2.

According to Table 1, the linear regression equation of the tomato's measured perimeter in pixels at maximum transverse diameter on the edge of 0° image is defined as

$$P = 0.2492\,n + 8.4294$$

in which P=tomato's estimated perimeter at maximum transverse diameter; n=pixels on the edge of 0′ image; and their correlation coefficient R=0.9990. The relationship obtained from experiments of the tomato's actual perimeter to estimated perimeter at maximum transverse diameter is shown in Table 3.

Table 1. Quantitative relationship between measured perimeter at maximal diameter and pixel in plane coordinate system of image

Sample	1	2	3	4	5	6	7
Pixels	928	787	667	861	806	937	767
Actual perimeter (mm)	240.19	203.59	175.90	223.10	210.41	241.51	200.23

Table 2. Quantitative relationship of tomato's area at maximum transverse diameter to pixels in plane coordinate system of image

Sample	1	2	3	4	5	6	7
Pixel	61918	44543	32127	53121	46822	62584	42579
Actual area (mm^2)	4590.94	3298.43	2462.13	3960.87	3523.02	4641.52	3190.41

Table 3. Relationship of tomato's actual perimeter to that of evaluated perimeter at maximum transverse diameter

Sample	Actual perimeter(mm)	Estimated perimeter(mm)	Absolute error(mm)	Relative error (%)
	197.31	199.57	2.26	1.15
2	217.43	217.26	0.17	0.08
3	173.13	173.65	0.52	0.30
4	221.59	220.25	1.34	0.60
5	237.08	234.95	2.13	0.90
6	218.95	217.26	1.69	0.77
7	245.32	244.92	0.40	0.16
8	247.74	244.67	3.07	1.24
9	194.17	193.09	1.08	0.56

Table 4. Relationship of actual area to evaluated area at maximum transverse diameter

Sample	Actual area (mm^2)	Estimated area (mm^2)	Absolute error (mm^2)	Relative error (%)
1	3098.03	3142.41	44.38	1.43
2	3762.08	3764.49	2.41	0.06
3	2385.34	2365.10	20.24	0.85
4	3907.50	3881.47	26.03	0.67
5	4472.82	4335.38	137.44	3.07
6	3814.99	3751.77	63.22	1.66
7	4789.09	4732.63	56.46	1.18
8	4884.31	4775.91	108.40	2.22
9	3000.11	2961.22	38.89	1.30

Seen from Table 3, the maximal error of the test perimeter exists within 3.07 mm, and the estimated perimeter at maximum transverse diameter gained by this means is of higher precision, which meets the

demand for identifying tomatoes.

According to Table 2, the linear regression equation of the tomato's estimated area in pixels at maximum transverse diameter inside the edge of 0° image is defined as

$$A = 0.0719n + 130.09$$

in which A=estimated area at maximum transverse diameter; n=pixels inside the edge of 0° image; and their correlation coefficient R=0.9995. The relationship obtained from experiments of the tomato's actual area to estimated area at maximum transverse diameter is shown in Table 4.

Seen from Table 3, the maximal error of the test area exists within 137.44 mm^2, and the estimated area at the maximum transverse diameter gained by this means is of higher precision, meeting the demand for identifying tomatoes.

4.2 To identify deformed tomatoes

Characteristic parameters of variations of fruit's diameter W_1—W_{12}, after being pretreated, are taken as inputs of the characteristic identification network. The network outputs are divided into three states: normal, passable and abnormal, which, if the *thermometer* method is adopted, can be respectively denoted as: (0, 0, 1), (0, 1, 1), (1, 1, 1).

The selection of GA's control parameters are as follows: population scale N is 50, crossing probability P_c is 0.4, mutagenic probability P_m is 0.001, and network's terminative condition is $MSE \leq 0.05$.

Through repeated selection, crossing, and mutagenesis, the sum of nodes that was ultimately obtained, in the optimal implicative layer of the characteristic network, was 8. Thus the valid network configuration was 12-8-3, and related connective weight values were stored in weight files. Twenty tomatoes were chosen as trained samples among which both normal fruits and deformed ones each composed half. Another twenty tomatoes were chosen as test samples. Experimental results are shown in Table 5.

Table 5. Identification results of deformed tomatoes

Variations of fruit's diameter / Identification results	Normal	Deformed
Normal fruits	10	0
Deformed fruits	0	10

The experimental results shown above indicate that identification rate of deformed tomatoes derived from GA trained neural network reaches 100%. It's necessary to emphasize that deformed fruits that are collectively selected for training must be sufficiently representative and comprehensive.

5 Conclusion

In the paper, roundness character is extracted from the tomato's image to identify cavernous fruit; and the character of fruit shape variations is used to identify deformed fruit. The character of fruit shape variations is analyzed by artificial neural network trained by GA. Research results indicate that through analyses of tomatoes' shapes, fruits with physiological diseases can be absolutely identified.

Correspondence to:
Junlong Fang , Changli Zhang
59 Wood Street
Gongbin Road, Xiangfang District
Harbin, Heilongjiang 150030, China.
Telephone: 01186-451-5519-0147
Cellular phone: 01186-13359709975
E-mail: swanhaha@163.com

References

[1] Ying Yibin, Jing Hansong, Ma Junfu, Zhao Jun, Jiang Yiyuan. Application of Computer Vision Technology to Test the Size of Virgin Pears and Flaws of Fruit's Surface[J]. Transactions of the Chinese Society of Agricultural Engineering 1999;15(2):197-200.

[2] Pan Wei. Study on Automatic Identification and Classification of Agricultural Products Using Computer Vision ——Automatic Identification and Classification of Tomato [D]. Northeast Agricultural University, 2000.

[3] Sarkar.N, Wolfe.RR. Feature Extraction Techniques for Sorting Tomatoes by Computer Vision[J]. Transactions of the ASAE 1985;28(3):970-4.

[4] Sarkar.N, Wolfe.RR. Computer Vision Based System for Quality Separation of Fresh Market Tomatoes[J]. Transactions of the ASAE 1985;28(5):1714-8.

[5] Zhang Changli, Fang Junlong, Pan Wei. Research on Genetic Algorithm Trained Multi-layer Feed-orward Neural Network's Automatic Measuring to Tomato Maturity[J]. Transactions of the Chinese Society of Agricultural Engineering 2001;17(3):153-6.

[6] Sun Yanfeng, Wang Zhongtuo. Parallel Genetic Algorithm[J]. Systematic Engineering 1995;13(2):14—6.

[7] Chen Genshe, Zhen Xinhai. Study and Development of Genetic Algorithm[J]. Information and Control 1994;23(4):215-21.

[8] Liang Min, Sun Zhongkang. Quick Learning Algorithm and Simulation Study of Multi-layer Feed-forward Neural Network[J]. Systematic E and Electronic Technology 1993(9):47-57.

[9] Buckley., J.J; Hayashi, Y.; Fuzzy Genetic Algorithm and Applications[J]. Fuzzy Sets and Systems 1994 Vol. 61:129—136.

[10] Ni.B, Paulsen.M.R, Reid.JF. Corn Kernel Crown Shape Identification Using Image Processing[J]. Transactions of the ASAE 1997,40(3):833—838.

[11] Liu He, Wang Maohua. Research on Artificial Neural Network Expert System of Fruit Shape's Identification[J]. Transactions of the Chinese Society of Agricultural Egineering 1996;12(1):171—6.

Study of Weak Signal Detection Based on Second FFT and Chaotic Oscillator

Chongsheng Li

Research Institute of Diagnostics and Cybernetics, Xian Jiaotong University, Xian, Shaanxi 710049, China
cslee@mail.xjtu.edu.cn

Abstract: Due to the incapacity of conventional method, while detecting periodic signals buried in the noise and deciding accurate frequency, a novel method is presented. This method is an integration of second FFT, which is an additional DFT after PSD, and the chaotic oscillator. Second FFT can increase the detection ability of weak periodic signals and the chaotic oscillator can improve frequency precision. This method is simple, quick and convenient in hardware realization and instrument design. Furthermore, its effect is demonstrated by detecting a simulation signal and a communication signal. [Nature and Science. 2005;3(2):59-64].

Keywords: weak signal detection; second FFT; chaotic oscillator; intermittent chaotic motion

1 Introduction

Quickly detecting weak periodic signals is widely used in radar, communication, sonar, earthquake and industrial measurement. The capability of conventional detection method such as power spectral density (PSD) is limited when the signal is buried in the noise. According to the character of PSD analysis, the more the data in PSD, the more powerful the ability of detection [1]. But it isn't practical, especially when the detecting speed is strictly required as online or real-time measurement being needed. Second FFT (Fast Fourier Transform) can improve the detection ability of weak periodic signals, but the frequency resolution is so low that the frequency of the to-be-detected weak periodic signal can't be decided accurately. The chaotic oscillator is proved to be effective in weak periodic signal detection for whose immunity to the noise and sensitivity to certain periodic signal, however, some blindness is showed when deciding the frequency. The method of integrating second FFT and chaotic oscillator is presented in this paper to learn from other's strong points to offset one's weakness.

2 Principle of Second FFT

Jen-Yi Jong presented a method, CPLE (Coherent Phase Line Enhancer), to discover periodic signals buried in noises utilizing the phase-coherent character among signals. The method restrains the disadvantage of phase information abandoned in conventional PSD analysis [2,3]. In order to run the method in a DSP (Digital Signal Processors) with high speed, we modified the method with not overlapping as segmenting data into blocks, which increases the calculating speed. We name the above method without overlapping second FFT and SFFT in short.

Assuming $x(n)$ is a discrete time series from a real signal after being sampled, segment $x(n)$ into M blocks with N data in every block, namely:

$x_1=\{x(1),x(2), \ldots,x(N)\}$

……

$x_M=\{x((M\text{-}1)\times N+1), \ldots,x(M\times N)\}$

The DFT (Discrete Fourier Transform) for a discrete time series x_i, is defined as:

$$X_i(k)=\sum_{n=0}^{N-1}x_i(n)e^{-j\frac{2\pi}{N}nk} \tag{1}$$

Make the following definition as:

$$X_i(k)=a_{ik}+j\cdot b_{ik} \tag{2}$$

where k=1,2, …,N, and i=1,2,…M.

From equation (2), we can construct a matrix Y with $M\times N$ form as shown in equation (3).

$$Y=\begin{bmatrix} a_{11}+j\times b_{11}\cdots a_{1N}+j\times b_{1N} \\ \cdots \quad \cdots \quad \cdots \\ a_{M1}+j\times b_{M1}\cdots a_{MN}+j\times b_{MN} \end{bmatrix} \tag{3}$$

namely: $Y=(y_1,y_2\ldots,y_M)^T$ and $y_i=X_i(k)$.

2.1 Conventional PSD

Conventional PSD calculates the mean value of the energy with respect to the frequency component according to equation (4), and then the value will be the estimation power of the corresponding frequency component. At last, power spectrum is obtained.

$$P_{CPSD}(k)=\frac{1}{M}\sum_{i=1}^{M}\left|\frac{1}{N}X_i(k)\right|^2 \tag{4}$$

2.2 SFFT

Conventional PSD will end after getting the power spectrum from equation (4), but SFFT will further deal with matrix Y. For no overlapping in segmenting the data, some middle processes are passed over. For whom are interested to this, please refer [2] and [3]. Carry out DFT in every column of matrix Y. If M is not a power of 2, we should pad with zeros. When M is the power of 2, $M_{new}=M$. When M is not the power of 2, M_{new} is the closest power of 2 larger than M. And then a new matrix R is obtained via carrying out an additional DFT along the column of matrix Y as shown in equation (5).

$$R=\begin{bmatrix} a'_{11}+jb'_{11} & \cdots & a'_{1N}+jb'_{1N} \\ \cdots & \cdots & \cdots \\ a'_{M_{new}1}+jb'_{M_{new}1} & \cdots & a'_{M_{new}N}+jb'_{M_{new}N} \end{bmatrix} \tag{5}$$

Get the max value of energy in each column and the corresponding index of w_{max} in matrix R, and calculate the power in a window according to equation (6) as the corresponding power with respect to the frequency component, where the rectangular window with length of 5 is used. The result $P_{SFFT}(k)$ is defined as the SFFT spectra.

$$P_{SFFT}(k)=\frac{1}{L_h}\sum_{w=w_{max}-W}^{w=w_{max}+W}a\times\left|E_k(w)\right|^2 \tag{6}$$

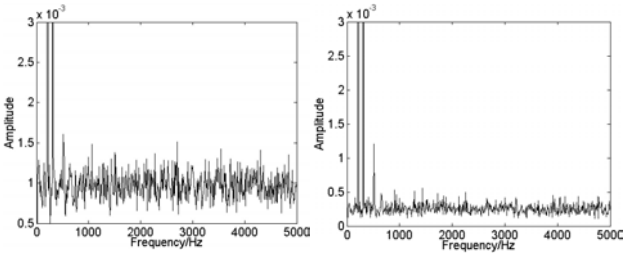
where $|E_k(w)|^2$ is the energy with respect to w_{max}; L_h is the length of window; W is the half length of window used in the method; α is amplitude correction constant and $\alpha=M_{new}/M$.

For example, the simulation signal $X(t)$ is defined as :

$$\begin{aligned} X(t)&=A_1\cdot\cos(2\pi f_1 t)+A_2\cdot\cos(2\pi f_2 t)\\ &\quad+A_3\cdot\cos(2\pi f_3 t)+5\times N(t)\end{aligned} \tag{7}$$

where t is time component; $N(t)$ is a Gaussian white noise with mean of 0 and variance of 1.

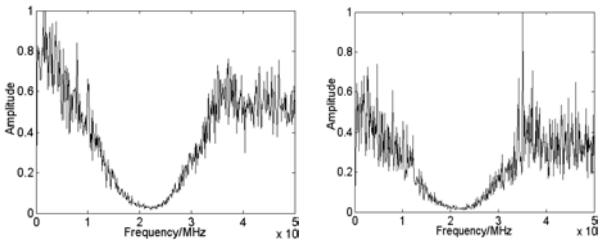
The sampling frequency is 10240Hz, A_1=10V, f_1=200.1Hz, A_2=1V, f_2=300.5Hz, A_3=0.3V, f_3=500Hz. Take 32768(=1024×32) data without overlapping, namely N=1024 and M=32. As illustrated in Figure 1(a), the amplitudes associated with frequency components of 200Hz and 300Hz are very obvious while 500Hz is not obvious. As illustrated in Figure 1(b), the amplitude of 500Hz is obviously clearer than in Figure 1(a).



(a) The conventional PSD normalized spectrum. **(b) The SFFT normalized spectrum.**

Figure 1. The comparison between conventional PSD and SFFT spectrum of simulation signal.

Utilizing SFFT, we deal with the BPSK (Binary Phase-Shifted Key) signal widely used in radar and communication. The sampling frequency is 100 MHz. In practice, the accurate carrier frequency is unknown, but some imprecise information is known. Subtract the mean value and then square BPSK signal; we can get the standard sine signal theoretically. Error-code will occur for sampling and signal-to-noise ratio (SNR) will be deteriorated for square, so general methods are incapable to detect the signal. Take 32768 (=1024×32) data without overlapping, namely N=1024 and M=32. As illustrated in Figure 2(a), the trajectory changes gently in entire frequency range and there is not very obvious peak. However, a peak is very obvious with respect to frequency about 35MHz, as shown in Figure 2(b).



(a) The conventional PSD normalized spectrum. **(b) The SFFT normalized spectrum.**

Figure 2. The comparison between the conventional PSD and SFFT spectrum of BPSK signal.

After lots of tests, it is demonstrated that SFFT is

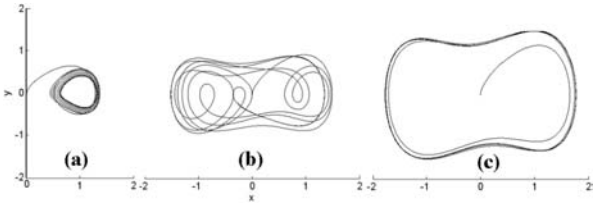


Figure 3. ω_0=1rad/s, c=0.5. (a) F_0=0.3 the small-scale periodic motion. (b) F_0=0.6 the chaotic motion. (c) F_0=0.9 the large-scale periodic motion.

more efficient than conventional PSD analysis on weak periodic signal detection in the same condition. Furthermore, performing a DFT on complex data is a well-known technique and the additional DFT will not increase the cost of hardware. However, the parameter N is limited to be small, for example, N=1024, due to the algorithm of DFT. So the frequency resolution, defined as F_s/N (F_s is the sampling frequency), is bigger and the frequency of weak periodic signal cannot be estimated accurately. As shown in Figure 1(b) and 2(b), we can only get the approximate frequencies of 200Hz, 300Hz, 500Hz and 35MHz.

3 Principle and Implement of Weak Periodic Signal Detection Using Chaotic Oscillator

Generally, for a nonlinear dynamic system, a small perturbation of system parameters may lead to the essential change of system state. Many methods utilizing the sensitivity to system parameters were put forward to detect weak periodic signal [4,5].

3.1 Principle

In this paper, the Holmes Duffing equation is chosen because it is one of the classic nonlinear systems and its characters have been studied extensively [6,7]. After introduction of the noise, the to-be-detected weak periodic signal and some transformations in time scale, the Holmes Duffing equation suit for any ω_0 is

$$\begin{cases} \dot{x} = \omega_0 y \\ \dot{y} = \omega_0(-cy + x - x^3 + F_0 \cdot \cos(\omega_0 t) \\ \quad + F_1 \cdot \cos(\omega_1 t + \psi) + N(t)) \end{cases} \quad (8)$$

where c is a damping constant; $F_0 \cdot \cos(\omega_0 t)$ is a periodic driving force in oscillator; $F_1 \cdot \cos(\omega_1 t + \psi)$ is the to-be-detected periodic signal. Here, $(F_1 \cos(\omega_1 t + \psi) + N(t))$ is named external perturbation.

If fix c, as F_0 varies gradually from zero to one threshold F_a, and then exceeds another threshold F_b, the oscillator state will vary from the small-scale periodic motion to the chaotic motion, and, at last, to the large-scale periodic motion in phase space. Discretize equation (8) and solve it using fourth-order Runge-Kutta algorithm. We chose iteration step h of 0.01s, total time of 50s, x_0=0 and y_0=0 convenient for investigating the state transformation of the Duffing oscillator as shown in Figure 3.

We can use the character of the phase plane trajectory of Duffing oscillator varying with F_0. Set ω_0 equal to the known frequency of the to-be-detected periodic signal and F_0 less than F_b slightly. The original state of oscillator is chaotic motion. When the periodic signal with the same frequency, namely $\omega_1=\omega_0$, is introduced, as long as $F_0+F_1>F_b$, even if F_1 is very small, the state transformation of the Duffing oscillator, from the chaotic motion to the large-scale periodic motion, will occur. By identifying the transformation, whether the periodic signal with frequency of ω_0 is present or not can be confirmed.

The threshold of Duffing oscillator can be obtained by Melnikov arithmetic [6,8] combined with experiments. As h=0.01 and ω_0=1rad/s, $F_b \approx 0.820$. Set F_0=0.815, F_1=0, and then add Gaussian white noise with mean of 0, variance of 0.5 and mean of 0, variance of 1 respectively. The corresponding phase plane trajectories are shown in Figure 4(a) and 4(b). Set F_1=0.01, ω_1=1rad/s, and then add Gaussian white noise with mean of 0 and variance of 1. The corresponding phase plane trajectory is shown in Figure 4(c). The phase plane trajectory fluctuates for being affected by noise, but the state of oscillator doesn't change. It is concluded that Duffing oscillator takes on some immunity to noise and strong sensitivity to some weak periodic signal.

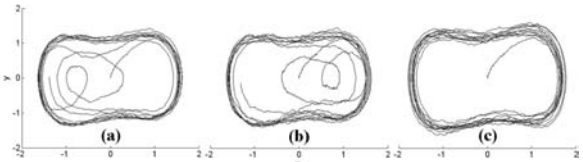


Figure 4. F_0=0.815, Duffing oscillator takes on some immunity to noise and strong sensitivity to some weak periodic signal. (a) σ^2=0.5, F_1=0, (b) σ^2=1, F_1=0 (c) σ^2=1, F_1=0.01

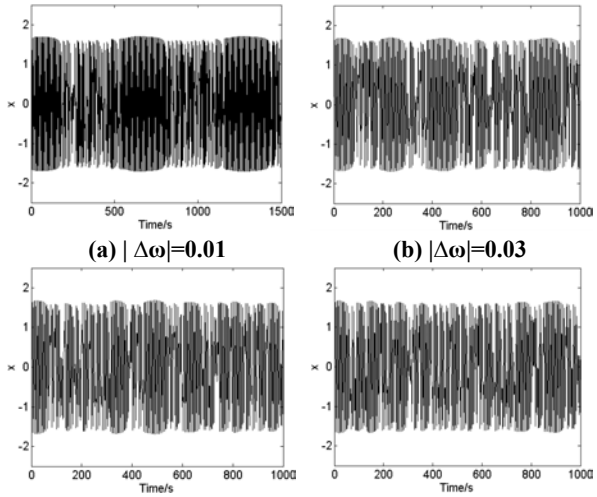
3.2 Intermittent Chaotic Motion

The situation when $|\Delta\omega|=|\omega_1-\omega_0|=0$, is analyzed above. Next, the situation with $|\Delta\omega|\neq 0$ is analyzed.

$$F(t) = F_0 \cdot \cos(\omega_0 t) + F_1 \cdot \cos(\omega_1 t + \psi)$$
$$= F_0 \cdot \cos(\omega_0 t) + F_1 \cdot \cos((\omega_0 + \Delta\omega)t + \psi)$$
$$F(t) = F_{CO}(t) \cdot \theta(t) \quad (9)$$
$$F_{CO}(t) = \sqrt{F_0^2 + F_1^2 + 2F_0F_1\cos(\Delta\omega t + \psi)} \quad (10)$$
$$\theta(t) = arctg\frac{F_1\sin(\Delta\omega t + \psi)}{F_0 + F_1\cos(\Delta\omega t + \psi)} \quad (11)$$

From equation (10), the total force $F_{CO}(t)$ will vary with $\Delta\omega$ and become a time function. As $F_0+F_1>F_b>F_0-F_1$, $F_{CO}(t)$ will be periodically more than or less than F_b with time. And then the periodic and chaotic motion of Duffing oscillator will happen alternately, namely intermittent chaotic motion. The period of alternate transform is just $T_1=2\pi/\Delta\omega$. Assuming $N(t)=0$, $\omega_0=1$rad/s, $F_0=0.815$, $F_1=0.02$, set $|\Delta\omega|=0.01$, $|\Delta\omega|=0.03$, $|\Delta\omega|=0.04$ and $|\Delta\omega|=0.05$ respectively, and the corresponding waveforms of oscillator's output in time domain are shown in Figure 5. When $\Delta\omega$ is very small, as shown in Figure 5 (a) and (b), theoretically T_1 is about 628s or 209s and $F_{CO}(t)$ varies more slowly than the state transformation, so the boundary between the chaotic and periodic motion is obvious and easy to be identified. However, when $\Delta\omega$ is relatively big, as shown in Figure 5 (c) and (d), theoretically T_1 is about 157s or 126s and $F_{CO}(t)$ varies relatively more quickly, so the boundary is not very obvious and hard to be identified. Therefore intermittent chaotic motion is limited by $\Delta\omega$. It's presented that $|\Delta\omega|\leq 0.03\omega_0$ is the range in which obvious intermittent chaotic motion happens [5].

Through the analysis and test above, we conclude that there is a serious limitation in weak periodic signal detection utilizing chaotic oscillator. The oscillator can only detect the periodic signal whose frequency is equal to inner periodic driving force or the frequency difference between them is in a small range. A parallel detecting array is presented to solve this problem in [5]. The detecting array is composed of 79 chaotic oscillators and the corresponding ω_0 is set from 1rad/s to 10rad/s with common ration of 1.03. When a signal is to be detected, transfer the frequency of signal into frequency range 1~10rad/s at first. And then introduce the signal to the detecting array. At last, the frequency of to-be-detected signal can be detected by identifying intermittent chaotic motion. However, when the frequency of to-be-detected signal is unknown, the scan of frequency is blindfold; any external perturbation will lead to false result and violent noise will weaken the detection ability. Furthermore, from Figure 5, hundreds, even thousands of seconds are needed to confirm the period of intermittent chaotic motion. It is slow and not suitable for quick detection. Furthermore, lots of elements are needed to construct 79 chaotic oscillators in hardware realization. The detection array needs great expense of hardware and its structure is complicated. The advantage of chaotic oscillator, easy to be realized with high speed in hardware, is missing.



(a) | Δω|=0.01 **(b) |Δω|=0.03**

(c) |Δω|=0.04 **(d) |Δω|=0.05**

Figure 5. The plots of intermittent chaotic motion.

4 Integration of SFFT and Chaotic Oscillator

To overcome the low frequency resolution of SFFT and utilize the property of intermittent chaotic motion, we present a new method based-on the integration of SFFT and the chaotic oscillator to detect weak periodic signals. First we use SFFT to find out all relatively obvious peaks in SFFT spectrum and estimate all frequency components associated with those peaks, and then set ±5% range of every estimated frequency to be the estimated frequency band and scan the frequency using a small array of Duffing oscillator in the frequency band. By this means, the accurate frequency can be obtained using only 5 chaotic oscillators.

4.1 Calculating Steps

(1). According to the number of relatively outstanding peaks in SFFT spectrum, record the estimated frequency values corresponding to the peaks. Get $f_e=(f_{e1}, f_{e2}, f_{e3}, \ldots)$, or $\omega_e=2\pi/f_e$.

(2). According to the sampling frequency and f_{e1}, adjust the playing speed of signal and interpolating or exampling the signal, then get a new sampling frequency f_{SN} and the post-treatment signal.

(3). According to f_{SN}, iteration step h and the new estimated frequency f'_{e1} are obtained.

(4). Let the post-treatment signal pass a band-pass filter with a $\pm 10\% f'_{e1}$ band in order to eliminate the disturbance of other frequency components.

(5). Take $\Omega=(0.95,0.975,1,1.025,1.05)$ ω'_{e1} into equation (8) respectively to replace ω_0, and get five equations. Get threshold F_b according to f_{SN} and iteration step h, and then set $F_0 \leq F_b$.

(6). Replace $(F_1\cos(\omega_1 t+\psi)+N(t))$ in above five equations respectively with post-treatment signal.

(7). By identifying the oscillator's output waveforms in time domain, the period of intermittent chaotic motion of every equation, and the four frequency difference, Δf_{e1}, Δf_{e2}, Δf_{e4} and Δf_{e5} are obtained. Take them into equation (12), and then the practical frequency f_{r1} of to-be-detected signal is obtained.

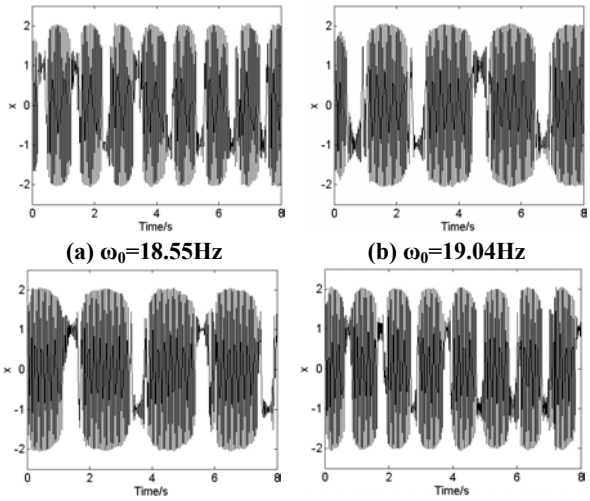
$$f_{r1} = c_{orf} \times ((0.95 f'_{e1} + \Delta f_{e1}) + (0.975 f'_{e1} + \Delta f_{e2}) + (1.025 f'_{e1} + \Delta f_{e4}) + (1.05 f'_{e1} + \Delta f_{e5})) / 4 \quad (12)$$

where c_{orf} is the frequency correction constant and relative to the playing speed and sampling frequency.

(8). Take ω_{e2} or f_{e2} and repeat step 2 to step 7, till all estimated frequencies have been worked out.

4.2 Application

Assuming the structure of simulation signal described in equation (7) is unknown. From Figure 1(b), we can see three obvious peaks, so the simulation signal is at least composed of three periodic signals and the corresponding frequencies are f_{e1}=200Hz, f_{e2}=300Hz, f_{e3}=500Hz. According to the estimated frequency f_{e1} and the sampling frequency of 10240Hz, we chose the playing speed of 1000Hz, so the iteration step h is 0.001s and c_{orf} is 10.24. Accordingly, the frequency 200Hz is adjusted to 19.53Hz. 19.53Hz is set as the centre of the estimated frequency band, and then 18.55Hz, 19.04Hz, 20.02Hz and 20.51Hz are other four frequencies. Take the five frequencies into equation (8) respectively, and get five equations. As mentioned above in step 6, we get Figure 6. As shown in Figure 6, intermittent chaotic motion happened in four oscillators. According to the adjacent degree between the frequency of the to-be-detected signal and the inner periodic driving force, the periods of intermittent chaotic motion in the four oscillators are different slightly. After calculating, $\Delta f_{e1}\approx 0.990$Hz, $\Delta f_{e2}\approx 0.501$Hz, $\Delta f_{e4}\approx -0.472$Hz and $\Delta f_{e5}\approx -1.010$Hz. From equation (12), f_{r1}=200.08Hz. Also f_{r2}=300.52Hz and f_{r3}=500.01Hz. It can be seen that error between the practical frequency and the frequency calculated from this method is extraordinarily tiny.

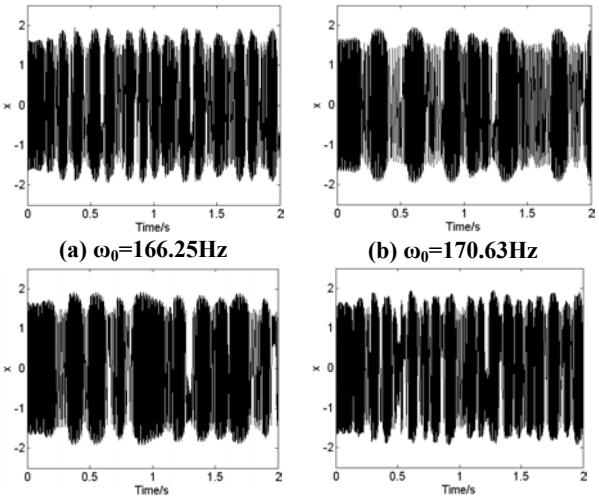


(a) ω_0=18.55Hz **(b) ω_0=19.04Hz**

(d) ω_0=20.02Hz **(e) ω_0=20.51 Hz**

Figure 6. The time domain waveforms of the four intermittent chaotic motions when simulation signal is introduced to the four Duffing oscillators respectively.

From Figure 2(b), the BPSK signal in section 2 is a single-frequency signal and the frequency is estimated about 35MHz, namely f_e=35MHz. According to sampling frequency of 100MHz and f_e, after pre-treating the signal, the iteration step h=0.0005 and the centre estimated frequency is 175Hz, other four frequencies are 166.25Hz, 170.63Hz, 179.38Hz and 183.75Hz. As mentioned above, the result is shown in Figure 7. Intermittent chaotic motions are not obvious and the corresponding periods are fluctuating for noise, so we adopt averaging method to get the periods of intermittent chaotic motion. After calculating, $\Delta f_{e1}\approx 8.68$Hz, $\Delta f_{e2}\approx 4.86$Hz, $\Delta f_{e4}\approx -5.42$Hz and $\Delta f_{e5}\approx -9.03$Hz. From above, f_r=34.955MHz with c_{orf}=200000.



(a) ω_0=166.25Hz (b) ω_0=170.63Hz

(c) ω_0=179.38Hz (d) ω_0=183.75Hz

Figure 7. The waveforms of the oscillator's output as BPSK signal introduced respectively.

The detection ability of weak periodic signals to be detected and frequency resolution are improved by means of SFFT and the chaotic oscillator. From figure 6 and figure 7, it is concluded that the larger the frequency of to-be-detected signal, the shorter the time required by calculation and the higher the detection speed. But this requires more powerful detection ability of chaotic oscillator. So we should choose the iteration step h and the threshold F_b according to the practical condition reasonably.

5 Conclusion

The method of integrating SFFT and chaotic oscillator presented in this paper is an integrated application of time domain and frequency domain method. SFFT is mostly used to estimate the frequency of the weak periodic signal approximately and the chaotic oscillator is used to decide the frequency accurately. Since a chaotic oscillator is composed of adder, integrator, gain element and sine signal generator, the detection system based on the chaotic oscillator is easy to be realized in circuits. This method improves not only the detection ability but also the frequency resolution of weak periodic signals, and the cost of detection system is not high and the method is easy in hardware realization and instrument design.

Correspondence to:
Chongsheng Li
Research Institute of Diagnostics and Cybernetics
School of Mechanical Engineering
Xian Jiaotong University
Xian, Shaanxi 710049, China
E-mail: **cslee@mail.xjtu.edu.cn**

References

[1] Alan V. Oppenheim, Alans S. Willsky, Iant T. Young, Signals and Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.

[2] Marshall Space Flight Center, Alabama, Acoustic-Emission-Analysis System for Diagnosis of Machinery, http://www.nasatech.com/Briefs/Aug00/MFS31468.html.

[3] Marshall Space Flight Center, Alabama, Coherent Phase Line Enhancer: a Method of Spectral Analysis, http://www.nasatech.com/Briefs/July01/MFS31426.html.

[4] Donald L. Birx. Chaotic Oscillators and Complex Mapping Feed Forward Networks (CMFFNS) For Signal Detection in Noisy Environments, IEEE International Joint Conference on Neural Network 1992;2:881~8.

[5] GuanyuWang, Dajun Chen, Jianya Lin, Xing Chen. The Application of Chaotic Oscillators to Weak Signal Detection, IEEE Transactions on Industrial Electronics 1999;46(2):440~4.

[6] John Guckenheimer, Philip Holmes. Nonlinear oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer-Verlag, New York, 1983:82~91.

[7] Kapitaniak, Tomasz. Chaos for Engineers: Theory, Application, and Control, Springer-Verlag, Berlin Heidelberg 2000:69~85.

[8] Liu Zengrong, Perturbation Criteria For Chaos, Shanghai Scientific and Technology Education Publishing House, Shanghai, China, 1994:36~67.

Workflow Timed Critical Path Optimization

Haibo Li[1], Dechen Zhan[2]

1 Centre of Intelligent Computing of Enterprises, School of Computer Science and Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150080, China
2 School of Engineer, Northeast Agriculture University, Heilongjiang 150030, China
Email: lihaibo@hit.edu.cn; dechen@hit.edu.cn

Abstract: Approaches to shorten workflow execution time have been discussed in many area of computer engineering such as parallel and distributed systems, a computer circuit, and PERT chart for project management. To optimize workflow model structure of workflow, an approach with corresponding algorithms is proposed to cut timed critical path of workflow schema, which has the longest average execution time path from the start activity to the end activity. Through systematically analyzing the dependency relationships between tasks at build-time, traditional optimization method of critical path is improved through adding selective and parallel control structures into workflow schemas. Data dependency rules are converted to control dependency rules according to semantic rules mined. Further more, consistency between tasks is guaranteed. Finally, to explain validity of the algorithm proposed, an experiment is provided to compare optimized model with original using critical path identification algorithm. [Nature and Science. 2005;3(2):65-74].

Keywords: workflow; critical path; data dependency relationship; control dependency relationship; control structure

1 Introduction

Workflow technology is an effective measure to change business processes in a more direct way. A workflow is the automation of a business process. In whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules (WMFC, 1995). To optimize workflow model and shorten execution duration of workflow is one of the most important way to improve efficiency of business processes. In order to shorten average execution times of workflow, the critical path in a workflow schema should be shorten first.

A simple definition of the critical path of a program is the longest, time-weighted sequence of events from the start of the program to its termination (Jeffrey, 1998). The critical path in a workflow schema is commonly defined as a path with the longest average execution time from the start activity to the end activity (Chang, 2002). The activities in the critical path are called critical activities. The execution times of activities in the critical path directly affect the total workflow completion time. The critical path has been widely discussed in many fields of computer engineering, e.g. evaluating the performance of large dynamic circuits (Lee, 1999), analyzing bottleneck of program in parallel and distributed systems (Jeffrey, 1998), determining the critical path in PERT chart for project management (Cormen, 1994). The concept of the critical path and the critical activity can be effectively utilized in many workflow issues, for example, workflow resource (Jin, 2001, Oh, 2000) and time management (Hagen, 1998; Pozewaunig, 1997; Heinl, 1998).

Many researches for the optimization algorithms of critical path (Gomory, 1961; Lenstra, 1977; Singh, 2000; Lam, 1977) were applied to parallel computing (Meajil, 2000; Hribar, 2001; Singh, 2001). Another approach to shorten critical path is used in project domain. For example, to check the rationality of schedule of every task in critical path, to make tasks on critical path parallel processing by decomposing them,

to support critical path by cutting resources on non-critical path, and to reschedule network structure are measures to shorten critical path (Hu, 1998). Traditional methods to analyze critical path in project domain are PERT (Program Evaluation and Review Technique) and CPM (Critical Path Method). But it is more complicated to shorten a critical path in workflow schema than in PERT chart. Firstly, we cannot use the previous methods (e.g. PERT) to shorten the critical path in a workflow schema, because they cannot support the two common control structures in workflow schema, i.e. selective structure and iterative structure (Chang, 2002). Secondly, optimizing a workflow model must guarantee consistency of all data dependencies and control constraints. For example, as response to change request, all data dependencies and control constraints between activities must be detected whether the problem of missing input or output values or cyclic waits may occur in the modified schema graph. Jin (2005) extended description power of PERT chart. Instead analyzing from workflow model domain, Jin (2005) used PERT chart to shorten workflow critical path as an assistant method. Other researches for the identification of critical path (Aalst, 1998; Cormen, 1994; Jin, 2005; Jin, 2001) only proposed a method to identify critical path in the context of a workflow, however, not gave an approach to shorten it.

To shorten the critical path in a workflow schema and meet all conditions when restructuring workflow graph, we must analyze data dependency systematically and control dependency between activities in workflow at first. We argue that the data dependency can be converted to data dependency between activities. Based on this convertibility, it is possible to shorten average execution time of critical path in a workflow schema by adding selective structure and iterative structure.

The remainder of the paper is organized as follows: In Section 2, we analyze data dependency and control dependency between activities in workflow schema and give some relative definitions. Section 3 gives some basic algorithms to guarantee consistency of workflow schema when changing. Section 4 presents our proposed method that systematically optimizes critical path in workflow schema. Section 5 gives experimental results to compare optimized model with original using critical path identification algorithm. Finally, we discuss the further work in Section 6.

2 Dependency Rule between Tasks

In workflow model, a task corresponds to a generic piece of work. A task is not defined for a specific business object, but for a type of objects, i.e., a task may be executed for many business objects. A business process is composed of one or more tasks following a certain order. An activity is one execution of a task. Each task in process is not isolated. The most obvious relationship between activities is logistic order which corresponds to a kind of partial order. This relationship is also called control dependency. These logistic relationships compose the control structures of the workflow. Four kinds of basic control structures, that is sequential, parallel (AND-Split, AND-Join), selective (OR-Split, OR-Join) and iterative (LOOP) structures, are defined in the workflow reference model (WDMC, 1995). Their semantics excerpted from WfMC (1999) are as follows.

(I) **Sequence:** Activities are executed in order under a single thread of execution, which means that the succeeding activity cannot start until the preceding activity is completed.

(II) **AND-Split:** A single thread of control splits into two or more threads which are executed in parallel within the workflow, allowing multiple activities to be executed simultaneously.

(III) **AND-Join:** Two or more parallel executing activities converge into a single common thread of control.

(IV) **OR-Split:** A single thread of control makes a decision upon which branch to take when encountered with multiple alternative workflow branches.

(V) **OR-Join:** Two or more alternative workflow branches re-converge to a single common activity as the next step within the workflow. No synchronization is required because of no parallel activity execution.

(VI) **LOOP:** A workflow cycle involves the repetitive execution of one (or more) workflow activities until a condition is met.

Except for control dependency between tasks, there exists data dependency relationship between tasks which expresses data access rule. All input data must be supplied before executing an activity, and after its successful completion all output data are written correctly. All data dependencies rules between activities should work and detect whether the problem of missing input or output values or cyclic waits may

occur at runtime. So data dependency rule should be well defined in workflow model to restrict the data access in workflow systems. In Figure 1, there exists control dependency (see solid lines) and its rules (predication P4, P5) between tasks T1, T2 and T3. There also exists data dependency (see dot lines which denote input and output relationship) and its rules (predication P1, P2, P3) among tasks T1, T2 and T3. The data dependencies for task T3 are checked. T3 is executable if all its data dependency rules are met. Value d3 produced by T1 must meet rule P3 before d5 is operated (read/written). Value d1 produced by other tasks must satisfy P1 before d4 is operated, and d2 is the same as d1.

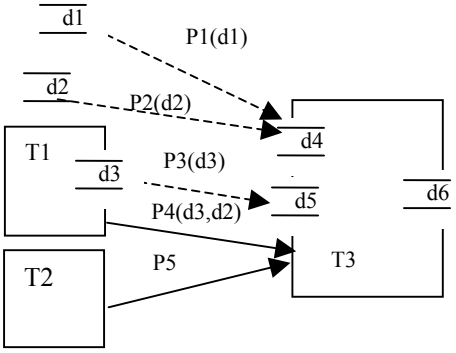


Figure 1. Dependencies relationship between tasks

The automation of business processes needs to be abstracted using a language, namely workflow specification language. The result is called workflow specification which contains information formally describing various aspects of a workflow (Chan, 1997). Considering dependency rules analyzed above, a workflow specification can be defined as follows:

Definition 1 (*Workflow specification*). A workflow specification, ws, is abstracted as a 4-tuple $< TN,CN,D,R >$, where

(i) $TN=\{t_1,t_2,...,t_n\}$ is a set of task nodes.

(ii) $CN=\{cn_1,cn_2,...,cn_n\}$ is a set of control nodes. Each element in CN has one of the above 6 types, that is Sequence, AND-Split, AND-Join, OR-Split, OR-Join or LOOP.

(iii) D is a set of data, used for tasks' input or output.

(iv) R is a superset of rules, $R=\{DR,CR\}$, DR is a set of data dependency rules, and CR is a set of control dependency rules.

Data dependency rule and control dependency role can be defined as follows.

Definition 2 (*Data dependency rule*). Let $t_i.d,..., t_k.d$ and $t_j.d$ be output data of tasks $t_i,...,t_k$ and input data of task t_j respectively, where $t_m.d \in D$. If $t_i.d,..., t_k.d$ need to be read before $t_j.d$ is created, say that there exists data dependency between $t_j.d$ and $t_i.d,..., t_k.d$. If condition $P_j(t_i.d,..., t_k.d)$=TRUE must be met before $t_j.d$ is created, say that P_j is a data dependency rule between data $t_j.d$ and $t_i.d,..., t_k.d$, denoted as $P_j(t_i.d,..., t_k.d) \rightarrow t_j.d$, where $P_j \in$ DR is the rule predication.

Definition 3 (*Control dependency rule*). For $\forall t_i,t_j \in$ TN, under the control of four basic control structures, if a partial order $t_i \prec t_j$ or $t_j \prec t_i$ between t_i and t_j can be determined, say there exists control dependency relationship between t_i and t_j, denoted as $t_i \xrightarrow{P} t_j$, where P is control dependency rule, for short, control rule.

Data dependency relationship between tasks can be classified as semantic and non-semantic rules. Non-semantic rule involves the validity of data chiefly, for example, to validate data format in database table. Semantic rule contains information which can be used to impact the course of business processes. For example, equipment management system in manufacture enterprise will execute different business branches according to different fault codes which origin from different equipment fault type. Due to lacking semantic analysis of business data, though structure of model is correct, data dependency relationships are always neglected. Only those semantic data dependency rules could be converted to control rules, so that we only discuss this sort of rule. Business rules can be embodied in software systems more exactly by mining data relationship.

3 Restructuring Model Schema

To achieve the goal of optimizing critical path of workflow, workflow model structure must be changed, so that execution time of critical path can be cut as short as possible. But traditional methods to analyze critical path such as PERT chart, does not support usual control structures: selective and iterative structure. More over, restructuring workflow schema must guarantee consistency of all data dependencies and control constraints. If task t2 has not completed while task t is ready to execute, in Figure 2 (a), data d2 may be

invalid, so task t and t2 cannot execute synchronously. In order to shorten critical path of workflow by add selective and parallel structures, consistency of all input and output data between tasks must be considered.

3.1 Background about critical path method

Numerous natural physical and organic processes exhibit behavior that is probably meaningfully modeled by Poisson processes. An important application of the Poisson distribution arises in connection with the occurrence of events of a particular type over time. The exponential distribution is frequently used as a model for the distribution of times between the occurrences of successive events such as customers arriving at a service facility. Because of them, the Poisson process and the exponential distribution have been used to analyze many areas of computer engineering. Hence, in this paper it is reasonable to consider workflow schema as a M/M/1 queuing network, as presented, see Chang (2002). We compare optimized model with original using critical path identification algorithm given by Chang (2002).

A workflow schema is represented with a set of nodes and directed edges. Tasks are interconnected by the six types of control structures. In a M/M/1 workflow queuing network, each activity is an independent M/M/1 queuing system at runtime of a workflow. Therefore, the arrival and departure rate in each activity can be specified, as well as the initial request rate to the start node, the service rate in each activity, and the branch selection-probabilities in each workflow control structure. So the average execution time of each task in a workflow schema can be determined. Consequently the path having longest average execution time in a workflow schema can be determined.

The execution time of each control structure is computed according to the following formulas:

(I) Sequence control structure: $W=\sum(1/\mu_i-\lambda_i)$, where λ_i is the arrival rate, μ_i is the service rate of task t_i.

(II) Selective control structure: $W=MAX(\sum(1/\mu_i-p_i\lambda_i))$, where $\sum p_i=1$, p_i is arrival probability of a branch.

(III) Parallel control structure: $W=MAX(\sum(1/\mu_i-\lambda_i))$.

Iterative control structure: $W=(1/p\mu_1-\lambda)+(1/p\mu_2-\lambda)+\ldots+(1/p\mu_n-\lambda)$

3.2 Consistency analysis

The correctness of workflow schema can be guaranteed by some rules, such as soundness property, which states that starting from the initial task node, it is always possible to reach any task nodes in a workflow schema, and for each reachable nodes, it is always possible to reach the finial task node (Aalst, 1998). Considering the six control types of workflow schema, for a split control structure, there must be a join control structure that can be combined with the split control structure whose resulting workflow schema forms a meaningful flow of control. From the meaning of the workflow control structures, when an AND-Split combines with an AND-Join and an OR-Split with an OR-Join, a workflow can be said to have a correct control structure. As a result, a workflow is an activity network in which activities are interconnected by workflow control where some control structures may be contained in some other control structures. In this paper we consider workflows meeting the following rules.

Rule 1: An AND-Split control structure should have its matching AND-Join control structure.

Rule 2: An OR-Split control structure should have its matching OR-Join control structure.

Rule 3: A non-sequential control structure can be completely contained in another non-sequential control structure, but two non-sequential control structures should not be partially overlapped. The rule is also called *nesting rule*. The outermost non-sequential control structure in a workflow is called a *control block*. The workflow depicted in Figure 4 includes a OR control block from node t1 to t7 and a LOOP control block. A control block can be nested in other control block. A control block starting from control node cn can be denoted as $Block(cn)=\{t_1,\ldots,t_m\}$, where t1,…,tm is activities involved in cn.

When logical orders between two task nodes are being changed, in principle, each task receives its input data from those tasks which already have been completed and produces its output data correctly. Data conflict must be avoided as well, i.e. two or more parallel tasks writing the same data item can lead to data conflict. Conflict problem are beyond this article's scope and can be seen in Minkyu (2001). So the following rules are given:

Rule 4 : For $\forall t\in TN$, and task set $T=\{t_1,t_2,\ldots,t_k\}$ which satisfies data dependency $P(t_1.d,\ldots,t_k.d)\rightarrow t.d$, where $t\notin T$, all tasks in T must be completed before t starts.

Rule 5 : Tasks in each branches of a parallel control structure can not write same data item. Let $Out(t_i)$ denote output data set of task t_i, so the rule can

also be described as following: $\cap Out(t_i)= \varnothing, i=1,\ldots,m, t_i \in Block(cn) = \{t_1,\ldots,t_m\}$.

When restructuring workflow graph, the five rules above must be held. The following sections give relevant restructuring steps.

3.3 Parallel control structure

When restructuring a workflow schema, to change sequential order of two tasks in critical order to parallel can cut average execution time. If this modification happens in a parallel control block, a second branch is only created, and if happens in a selective or iterative control block, a new parallel control block is nested in the control block. Adding more semantic rules can improve flexibility of business processes. These semantic rules come from business rules. So new semantic rules added must guarantee consistency of data dependency rule between each task. The following algorithm for restructuring to parallel control structure is given.

First, let TB_{i-1} and TB_i be two tasks or control blocks in path, which is denoted as path=<cn_{i-1}, TB_{i-1}, cn_i, TB_i, cn_{i+1}>. Predecessor(n) and Successor(n) denote predecessor and successor of node n respectively, where $n \in TN$ or $n \in CN$. Type(cn) and Type(Block(cn)) $\in$ {SEQ, AND-Split, AND-Join, AND-Split, AND-Join, LOOP } denote type of cn and Block(cn) respectively. Let type(cn_i)=SEQ where cn_i is involved in the path.

Algorithm 1 CreateParallel (TB_{i-1}, TB_i)
Input:TB_{i-1}, TB_i
Output: control block b
Begin
b:=∅
Predecessor(TB_{i-1}) := cn
Predecessor(TB_i) := cn
Type(cn):= AND_Split
Successor (TBi-1) := cn'
Successor (TBi) := cn'
Type(cn'):= AND_Join
b := Block(cn)
return b
End

3.4 Selective control structure

To change sequential order of two tasks in critical order to selective can still cut average execution time. If this modification happens in a selective control block, a second branch is only created, and if happens in a parallel or iterative control block, a new selective control block is nested in the control block. Changing to selective control structure must keep consistency of data dependency rule between tasks also. The algorithm given below is to restructure selective control structure.

Algorithm 1 CreateParallel (TB_{i-1}, TB_i)
Input:TB_{i-1}, TB_i
Output: control block b
Begin
T':= ∅
Predecessor(TBi) :=cn
Predecessor(T'):=cn
Type(cn):= OR_Split
cn':= OR_Join
Successor (TB_i)) := cn'
Successor (TB_{i-1}) := cn'
Type(Successor (TB_{i+1})) := SEQ
Type(Successor (T')) := SEQ
b := Block(cn)
return b

4 Critical Path Optimization Method

Through mining data dependency relationship between tasks, we can change workflow control structures and shorten the critical path in a workflow schema until the execution time of the path can not be cut any more. The resulting schema is optimum under effect of current set of data dependency rules. The more semantic rules are mined, the more probability to optimize workflow schema by restructuring it. Based on the critical path theory mentioned in section 3.1 , changing more tasks without data dependency in workflow schema to selective or parallel structure can shorten execution time of the critical path.

Let DR be a set of data dependency rules, $D=\{D_1,\ldots,D_n\}$ be set of all tasks' data where $D_i \in D$ is output data set of task T_i. Algorithm 3 begins to deal with all data dependency $P_{ij}(d_p,\ldots,d_q) \rightarrow T_i.d_{ij}$ in DR, where $d_{ij} \in D_i, Din_{ij}=\{d_p,\ldots,d_q\}$, $Din_{ij} \cap D_i=\varnothing$.

Algorithm 3 Optimization(T_i)
Input: Task node T_i
Output: a set of control rule CR
Begin
While DR≠∅
For $P_{ij} \in$ DR do

If $D_{i-1} \cap Din_{ij} = \varnothing$ and $Out(t_{i-1}) \cap Out(t_i) = \varnothing$ then

CreateParallel(T_{i-1}, T_i)

else

CreateAlternative (T_{i-1}, T_i)

End for

CR:= CR $\cup$ { P_{ij} }

DR:= DR- { P_{ij} }

End while

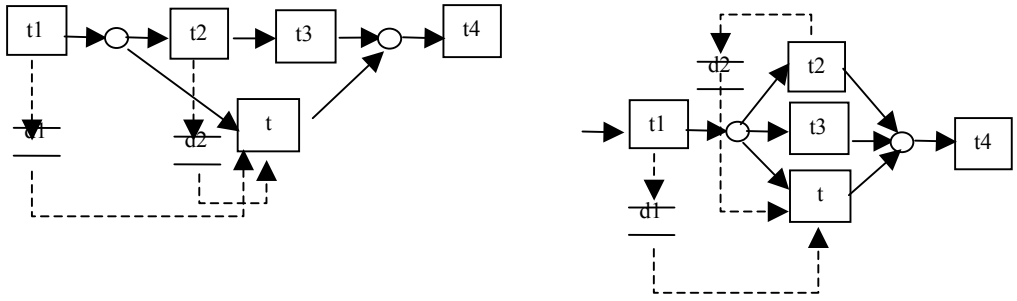
Return DR

End

The statement "if" in algorithm 3 above judges data dependency and data set of written between two neighbouring tasks, which can guarantee satisfying rule 4 and 5. The algorithm can also deal with two neighbouring control blocks to validate data dependency, in which circumstances the granularity is enlarged to control block from task so that control blocks can be combined, see algorithm 1 and 2.

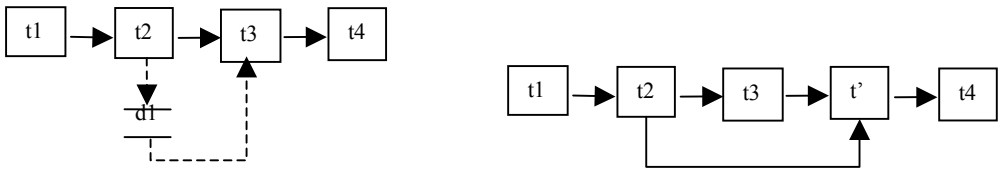
A parallel control structure can be created because no any data dependency between task t2 and t3 in Figure 2 (a) (we adopt expression of workflow schema, see Chang(2002), here we don not give unnecessary details). Data item d1and d2 are input data of task t. d2 may be an invalid data when t begins to execute, therefore restructuring to Figure 2(b) can not happen for not meeting rule 4. In Figure 3(a), there exists data dependency between task t2 and t3, which can be converted to control rule, so a selective branch is added and a new control block is created. It is noted that null task node t' is used to "shortcut" t3. This treatment method can be compatible with the circumstances of nested by other control blocks and meet rule 1, 2 and 3. The example in section 5 shows the soundness of control structure.

After computed by algorithm 3, workflow schema restructured shows a new partial order of different tasks. However the new business logic maybe doesn't match with business custom of enterprises. So some partial orders should be kept artificially when executing algorithm 3.



（a） （b）

Figure 2. Creating parallel control structure



（a） （b）

Figure 3. Creating selective control structure

5 Application and Example

The proposed approach has been applied to the development of our CERP system for a large number of different manufacture enterprises in some cities of China. In development phase of ERP software system and at run time, there are more possibilities of restructuring a workflow schema along with more semantic rules found by domain experts or developers. In order to validate our approach, we give equipment repair processes as an example in most manufacture enterprises. The tasks involved in the process and their meanings are seen in table 1. The content of business process is introduced as following. When some equipment does not work, a requisition is summated from workshop to Department of Equipment Management (DEM). For a small fault, maintenance man of workshop repairs it by himself. If he cannot fix up the machine, another requisition needs to be summated to DEM again and the engineer of DEM completes the business process of repair which involves assigning task, drawing part or dissipative material and repairing. If the failure part is found not to be resumed, another requisition needs to be summated to DEM once more. The repair task is relegated to others out of the enterprise.

Depending on our experience, we assume that the service requests arrive with rate $\lambda=4$. There are three different paths in the workflow Figure 4 when we do not consider outmost iteration structure. Let the sequence of tasks t1, t2, t7, t8 be called *path 1*, the sequence of task t1, t3, t5, t6, t7 and t8 be called *path 2*, the sequence of task t1, t4, t7 and t8 be called *path 3,* and probability of each branch of the selective structure which locates between t1 and t2, t3, t4 is $\alpha1=0.4$, $\alpha2=0.4$ and $\alpha3=0.2$ respectively. Service rate μ of each task is given in Table 1.

Table 1. Service rates ($\mu=1/t$ tis service time)

Tasks	Content	Service Rate μ(s)
t1	Summit requisition for repairing	10
t2	Draw part or material	5
t3	Assign task for dispatching person	10
t4	Relegate repair to other enterprise	8
t5	Assign task for computing part or material	8
t6	Draw part or material from storage	6
t7	Fix up ,check and accept	10
t8	Collect data of repair	8

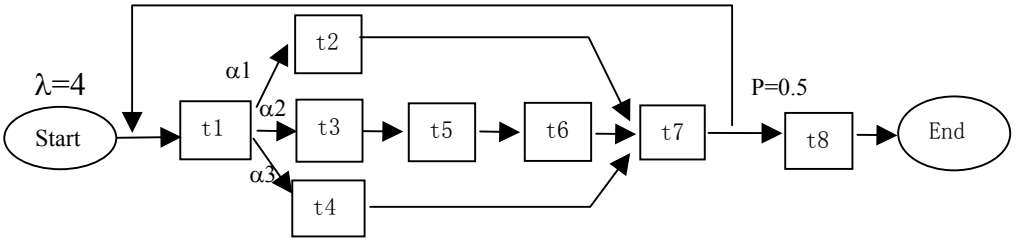


Figure 4. Business process of repair a equipment

The following is the process of computing critical path.

Step 1: Compute average execution time of each branch of the selective structure.

$W_2=1/\mu_2-\lambda_2=1/5-\alpha_1\times\lambda=0.294$s

$W_{356}=1/(\mu_3-\lambda_3)+1/(\mu_5-\lambda_5)+1/(\mu_6-\lambda_6)=1/(10-\alpha_2\times\lambda)+1/(8-\alpha_2\times\lambda)+1/(6-\alpha_2\times\lambda)=0.502$s

$W_4=1/\mu_4-\lambda_4=1/8-\alpha_3\times\lambda=0.139$s

The maximum is $W_{356,}$, then the path with longest average execution time in the selective structure is t3, t5 and t6.

Step 2: Compute average execution time of the iterative structure.

$W_1=1/p\mu_1-\lambda_1=1$; $W_7=1/p\mu_7-\lambda_7=1$

$W_{13567}=W_1+1/(p\mu_3-\alpha_2\times\lambda)+1/(p\mu_5-\alpha_2\times\lambda)+1/(p\mu_6-\alpha_2\times\lambda)+W_7=3.425$s

Step 3: Compute average execution time of the whole process.

$W_8=1/\mu_8-\lambda_8=1/8-4=0.25$s, because task t8 has already been in critical path.

The finial average execution time of critical path is: $W=W_{13567}+W_8=2.752$s, so we get the critical path is t1, t3, t5, t6, t7 and t8.

Then, in order to optimize critical path, semantic rules should be mined first. According to our experience in some manufacture enterprises and analysis to process of equipment repair, we find the fact that some parts are repaired by being relegated to other company. The property of the part should be provided as soon as possible. So the well-timed position is when task t7 is completed for the first execution. The semantic rule provided are described as *IsNotConsign (t7.PartID)→t3.PartID.* A selective branch can be added after task t1 following algorithm 3. In addition, there is not any data dependency between task t3 and t5, so the sequential order can be converted to parallel. Figure 5 shows workflow schema after conversion.

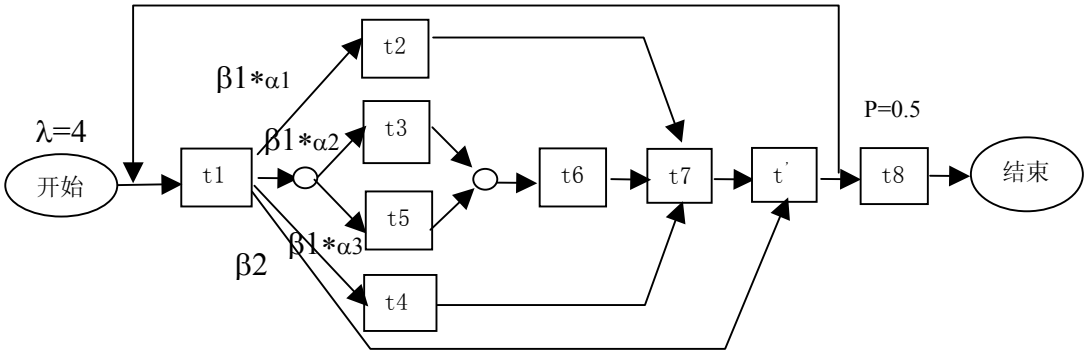


Figure 5. Workflow schema after converting

The following is the process of computing critical path. Let λ=4,α1=0.4, α2=0.4, α3=0.3,β1=0.5, β2=0.5.

Step 1: Computing the average execution time of the innermost parallel structure.

$W_{35}=\mathrm{MAX}(1/(\mu_3-\lambda_3),1/(\mu_5-\lambda_5))=\mathrm{MAX}(1/(\mu_3-\beta1*\alpha2*\lambda),1/(\mu_5-\beta1*\alpha2*\lambda))=0.139$s, task t5 is in critical path.

Step 2: Computing W_{56}.

$W_{56}=1/(\mu_5-\lambda_5)+1/(\mu_6-\lambda_6)=1/(\mu_5-\beta1*\alpha2*\lambda),1/(\mu_6-\beta1*\alpha2*\lambda)=0.331$s

Step 3: Computing W_{2564}.

$W_{2564}=\mathrm{MAX}(1/(\mu_2-\lambda_2),W_{56},1/(\mu_4-\lambda_4))=\mathrm{MAX}(1/(\mu_2-\beta1*\alpha1*\lambda),W_{56},1/(\mu_4-\beta1*\alpha3*\lambda))=0.331$s

Step 4: Computing $W_{1567t'}$

$W_{1567t}=1/(\mu_1-\lambda_1)+W_{56}+1/(\mu_7-\lambda_7))=1/(\mu_1-\beta1*\lambda)+W_{56}+1/(\mu_7-\beta1*\lambda))=0.581$s.

Here the execution time of null task t' is zero, then let $1/\mu\to\infty, W_{t'}\to0$

Step 5: Computing average execution time of iteration structure.

$W_{1567}=1/(p*\mu_1-\lambda)+1/(p*\mu_5-\beta1*\alpha2*\lambda)+1/(p*\mu_6-\beta1*\alpha2*\lambda)+1/(p*\mu_7-\beta1*\lambda)=2.1$s.

The finial average execution time of critical path is: $W'=W_{1567}+W_8=2.35$s, so we get the critical path is t1, t5, t6, t7, t', t8.

By comparing the workflow schema Figure 5 with Figure 4, in optimized critical path, because the order of

task t3 and t5 are changed to parallel structure, the average execution time begin to cut from t1 to t5 (here t3 does not belong to critical path, so its execution time is zero) in Figure 6. At the same time, adding selective branch can result in a reducing probability of all branches, so that the average execution times of t6 and t7 reduce accordingly.

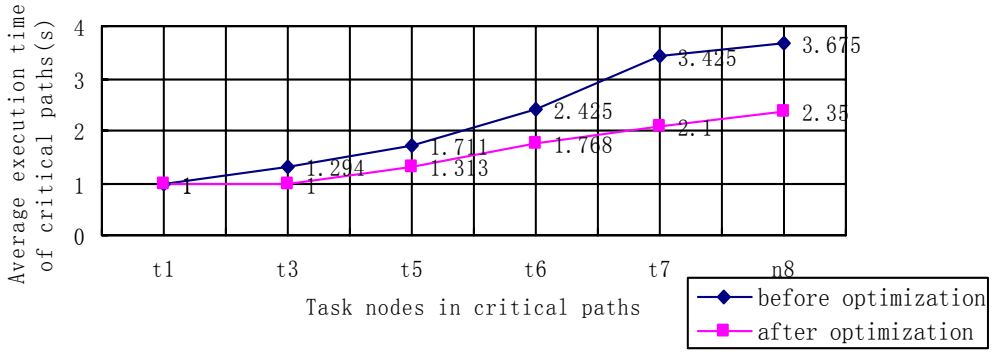


Figure 6. Two critical paths' execution times

6 Conclusion

The primary benefit of critical path method is to help domain experts or models of developers understand where they can most productively spend time modifying their models. However that is not enough only to locate the position of critical path because they still lack an effective approach to optimize it in former researches, such as PERT chat. Though some researchers attempt to optimize workflow model adopting the approach of PERT char, there are also some limitations to the technique in which selective and parallel control structure are not supported, as mentioned above. We show that it is operable to restructure a workflow schema by mine more semantic rules, i.e. data dependency rule called in this paper. The changed model is called optimization in current data dependency rule set.

The paper proposes an approach to optimize workflow construct by shortening average execution time of workflow critical path which has the longest average execution time from the start activity to the end activity. We improve traditional optimization method of critical path, and add selective and parallel control structures by mining more data dependency relationship between tasks. So the approach is also called a conversion from data dependency rule to control dependency.

In addition, this conversion should keep rationality of business process logically, conform to business custom of enterprises, and be handled under human intervention. It is dogmatic to convert all mined semantic rules in order to optimize workflow graph. Unreasonable business logic must result in unpractical business process which can be applied.

Acknowledgments

The authors also wish to acknowledge the financial support of the National High-Tech. R&D Program for CIMS, China, Grant 2003AA4Z3210, 2003AA413023 and 2002AA413310.and another cooperative project from European Union.

Correspondence to:
Haibo Li
Centre of Intelligent Computing of Enterprises
School of Computer Science and Engineering
Harbin Institute of Technology
Harbin, Heilongjiang 150080, China
Working: School of Engineer
Northeast Agriculture University
Harbin, Heilongjiang 150030, China
Telephone: 86-451-86412664
Email: lihaibo@hit.edu.cn

References

[1] Cormen TH, Leiserson CE, Rivest RL. Introduction to Algorithms, 1994 (MIT Press, Cambridge, MA).
[2] Duk-Ho Chang, Jin Hyun Son, Myoung Ho Kim. Critical pPath iIdentification in the cContext of a wWorkflow. Information and Software Technology 2002;44(7):405-17.
[3] Gomory RE, Hu TC. Multi-terminal network flows. SIAM 1961;9:551-70.

[4] Hagen C, Alonso G. Flexible exception handling in the oOpera process support system. Proceedings of the Eighteenth IEEE International Conference on Distributed Computing Systems. 1998.

[5] Heinl P. Exceptions during workflow execution. Proceedings of the Sixth International Conference on Extending Database Technology, workshop on workflow management, Valencia, Spain, 1998.

[6] Hollingsworth D. The workflow reference model. Workflow Management Coalition. 1995.

[7] Hollingsworth JK. Critical Path Profiling of Message Passing and Shared-Memory Programs. Transactions on Parallel and Distributed Systems 1998;9(10).

[8] Hribar MR, Taylor VE, Boyce DE. Implementing parallel shortest path for parallel transportation applications, Parallel Computing 2001;31:1537-68.

[9] Hu Yunquan. Operational research foundation and application (in Chinese). Harbin Institute of Technology Press , Harbin China 1998:147-8.

[10] Jeffrey K. Hollingsworth JK,. Member, IEEE Computer Society. Critical Path Profiling of Message Passing and Shared-Memory Programs. IEEE Transactions on Parallel and Distributed Systems. 1998;9(10):1029-40.

[11] Jin Hyun Son , Myoung Ho Kim. Analyzing the critical path for the well-formed workflow schema. Proceedings of the Seventh International Conference on Database Systems for Advanced Applications (DASFAA .01), 2001.

[12] Jin Hyun Son, Jung Sun Kim, Myoung Ho Kim. Extracting the Workflow Critical Path from the Extended Well-Formed Workflow Schema. Journal of Computer and System Sciences 2005;70(1):86-106.

[13] Lee KT, Abraham JA. Critical path identification and delay tests of dynamic circuits. Proceedings of International Test Conference. 1999:421-30.

[14] Lenstra JK, Rinnooy Kan AHG, Brucker P. Complexity of Machine Scheduling Problems. Ann Discrete Math 1977;1:343-62.

[15] Meajil A, El-Ghazawi T, Sterling T. Characterizing and representing workloads for parallel computer architectures. Journal of Systems Architecture. 2000;46(1).

[16] Minkyu Lee, Dongsoo Han, Jaeyong Shim. Set-based access conflict analysis of concurrent workflow definition. Information Processing Letters 2001;80:189–94.

[17] Oh SK, Son JH, Lee YJ, Kim MH. An efficient method for allocating workflow tasks to improve the performance of distributed workflows. International Conference on Computer Science and Informatics. 2000:445-8.

[18] Pozewaunig H, Eder J, Liebhart W. ePERT: Extending PERT fro workflow management systems. The First European Symposium in ADBIS 1997:217-24.

[19] Shui Lam, Ravi Sethi. Worst Case Analysis of Two Scheduling Algorithms. SIAM J Comput 1977;6(3):518-36.

[20] Singh G, Zinder Y. Worst-case performance of critical path type algorithms. International Transactions in Operational Research. 2000;7:383-99.

[21] Singh G. Performance of critical path type algorithms for scheduling on parallel processors, Operations Research Letters 2001;29:17-30.

[22] Son JH, Kim MH. Improving the performance of time-constrained workflow processing. Journal of Systems and Software 2001;58(3):211-9.

[23] van der Aalst WMP. The application of Petri nets to workflow management. Journal of Circuits, Systems and Computers 1998;8(1):21-66.

[24] Workflow Management Coalition: Process Definition Interchange, Document No. WfMC-TC-1016-P, 1999.

Studying on Flood Damage Assessment System of Hunhe River Basin

Jinping Zhang, Junshi He, Lina Cao

College of Water Resources, Shenyang Agricultural University, Shenyang, Liaoning 110161, China;
Telephone: 01186-13998386658; E-mail: hejunshi@163.com

Abstract: It is necessary to establish an evaluation system to calculate flood disaster losses quickly, accurately and wholly for effectively reducing or avoiding flood losses and promoting correctness and validity of flood control and regulation decision of Hunhe River. From the view of flood control and disaster reduction of Hunhe River, this paper designs flood damage evaluation system of Hunhe River basin and also describes its functions and characteristics. [Nature and Science. 2005;3(2):75-78].

Key words: Hunhe River; flood control and regulation decision; flood disaster evaluation

Introduction

Hunhe River is the larger river in the middle part of Liaoning province of China, and its drainage area is 11481 km^2, which is great important for economic development of Liaoning province. However, Hunhe River is also the region with frequent flood. In history, many extraordinary flood had ever occurred here, especially the flood in 1995, which caused a great damage for people's lives and properties. Inundated areas caused by the flood in1995 is 1356 km^2, and stricken population attains to 0.433 million, making up 62.6% of the total population in floodplain, and submerging farmland is 0.0946 million hm^2, making up 80.5% of the total farmland in floodplain. The huge flood damage losses not only bring about economic waste, but also greatly influence people's mentality health in stricken area. Therefore, evaluating system assessing flood losses quickly and accurately is urgently needed to assist flood control regulation decision, update flood pre-scheme and make full play of structural measures and non-structural measures in order to avoid or reduce flood losses. This system will promote the healthy and stable development of economy and society of the nation.

Aiming at the characteristics (namely: extensive influencing area, long consistent time and severe losses) of flood damage losses in Hunhe River, and stressing on the basis of flood losses mechanism, flood damage assessment system of Hunhe River is developed. This system can act as an effective sustentation for flood control decision support system of Hunhe River and service the whole process of flood control regulation decision.

1 Aim and Principles of System Development

Aim of system development is to offer a scientific, rational and reliable flood damage assessment system. The following principles should be insisted in system: （1） practicality, credibility, stability, expand; （2） having favorable user interface and mutual accomplish between human and computer; （3） the design of module must be rational, it not only guarantees the data transformation quickly and effectively among modules in system, but also ensures transformation of other modules; （4） self-governed database is owned and rationally accepted in system; （5） realizing effective flood damage assessment; （6） obligating fetches for other systems.

2 General Design

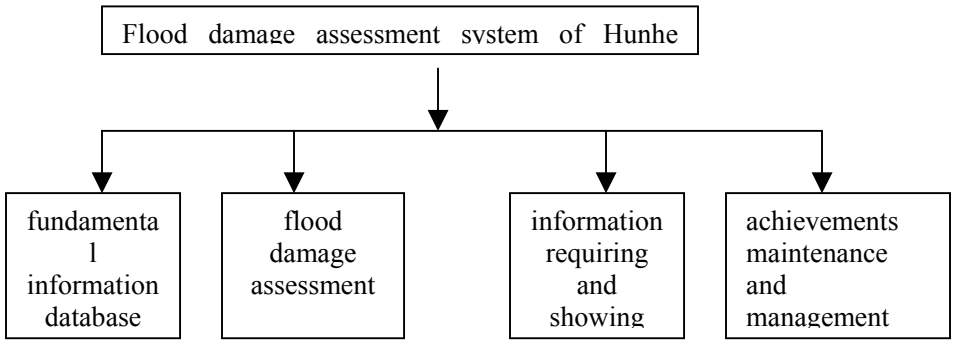


Figure 1. System configuration

2.1 System configuration

The system includes four parts: （1） fundamental information database; （2） flood damage assessment model; （3） information requiring and showing; （4） achievements maintenance and management (Figure 1).

2.2 System function

2.2.1 Fundamental database

There are settlement and properties in floodplain which is the reason why the flood can cause great damage, when flood occurs, it will do harm to human. So, flood disaster is the process in which flood acts on human's society under certain natural geography condition, such as natural factors, environmental factors and social factors. When studying on flood disaster, we must aim at these three factors seriously to achieve overall flood disaster assessment. Thus, fundamental data should be included in fundamental database: flood characteristic data （begin and end time, lasting time, submerging time, velocity of flood flow, historical flood damage losses and its occurring times, various losses data, etc）; geography information data in floodplain (topography, physiognomy, soil, vegetation, land utilization condition, administration district, river basin district, water project, etc); society and economy data (population, numbers of houses, farmland area, total area, total national economic production value, fixed production value, industrial production value, agricultural production value, transportation and infrastructure condition, prevalence degree of flood control and disaster reduction knowledge, etc).

For realizing dynastic interlinkage between user and data and for the convenience of update database timely, Access database is adopted to manage and maintain the whole data so as to provide with timely and accurately information for flood control decision.

2.2.2 Flood disaster losses assessment model

Flood disaster losses assessment model is a core of this system. Flood disaster losses assessment is to evaluate caused losses, possible losses and causing losses. Form this view, it includes three parts: pre-disaster assessment, mid-disaster assessment and sith-disaster assessment. Pre-disaster assessment and mid-disaster assessment belong to predictive assessment, while sith-disaster assessment is actual losses data. This system mainly aims at pre-disaster assessment and mid-disaster assessment, but because the influencing factors of flood disaster are various and complicated, and also have the dynamic and uncertain characteristics, it is impossible to evaluate damage losses caused by flood accurately. In a given region, on the basis of flood disaster losses mechanism, whether or not the flood losses are serious depends on magnitude and scale of flood itself. There are lots of indexes to weigh its magnitude and scale, but in out nation and aboard, many scholars accept the relationship between inundated water level and flood losses rate to predict flood losses. This is because the water level is higher and the losses are larger. So this relationship has fairly precision and reliability. In history, when a flood is recorded, it mainly referred to its flood water level, thus this system also

applies the relationship curve to pre-evaluate flood disaster losses.

The main content of relation curve of water level and losses rate is that according to information of history flood (such as inundated water level), social and economic information and other hydrologic information, flood disaster losses rate is calculated, namely: the ratio between inundation losses and original value, then we can establish relation curve of inundation water level and flood damage losses rate.

When user wants to compare different assets losses under the same water level, he can select the same water level as he inputs flood water level, while inputs various assets as selects assets sorts. At the same time, he must inputs according assets amounts, after clicking "predication" button, and then it can get various asset loss values under the same water level. To compare asset losses better, this system also set up column diagram. When user clicks "displaying" button, a column figure will be shown in the interface. Thus, an intuitionist knowledge is achieved (Figure 2).

Figure 2 - System configuration of model:

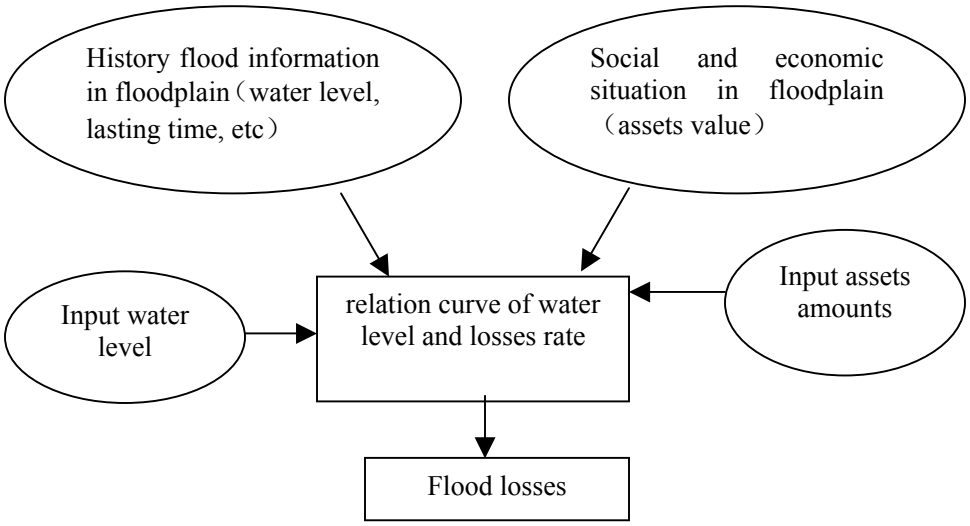


Figure 2. Model configuration of water level and losses rate

On the basis of inundation water level of various schemes and combing with social and economic information, flood damage assessment model predicts possible economic losses resulted from flood with various frequencies to assist to determine flood control standard and disaster reduction measures in a giver region. For cooperation with flood control planning of Hunhe River, direct economic losses in this paper include: industry, agriculture, transportation, business, postage, water engineering, family assets, etc. After inundation water level and corresponding loss rate are determined, we can multiple flood loss rate with assets value in direct economic losses, then to plus flood losses of various assets, flood damage direct economic losses in the whole studying area will be obtained.

2.2.3 Requiring and showing

Requiring and showing function can help flood control decision and management personnel to analyze flood disaster situation and query history flood, to inquire about needed information and related model parameter in pre-disaster, mid-disaster and sith-disaster, to predict and calculate flood water level and all asset loss values.

2.2.4 Achievement management

Achievement management mainly manages evaluation results of flood disaster losses, including the statistics and management of all kings of disaster information; report forms management and mimeograph; flood loss figures drawing of each item; editing, deleting, adding all sorts of evaluation achievement; etc.

3 Conclusion

Development of this system is in favor of flood control and disaster reduction decision in Hunhe River Bain. On the base of the quantitative relationships of characteristic (flood water level) of flood itself and social and economic losses, predicative flood loss evaluation caused by a flood is carried out. Meanwhile, for the convenience of real-time requiring, dynamic linkage is realized between data and model to effective dialogue between user and computer. So user directly participates in the whole operation process of model to reduce or avoid flood losses with more scientific and rational flood control and disaster reduction measures.

Correspondence to:
Junshi He
College of Water Resources
Shenyang Agricultural University
Shenyang, Liaoning 110161, China
Telephone: 01186-13998386658
E-mail: hejunshi@163.com

References

[1] Flood Control and Drought Relief Command Office in Liaoning Province. Extrodarnary Flood in 1995 in Liaoning Province. Shenyang. Liaoning Scientific and Technological Publishing Company. 1999.

[2] Wang Yanyan, at al. Development of Flood Damage Assessment System of Shanghai. Journal of Catastrophology 2001;16(2):7−12.

[3] Zhang Linpeng. Information Management System of Vulnerability for Evaluation of Flood Disaster. Journal of Natural Disasters 2002;11(4):66−72.

[4] Liu Jun, Xu Xiangyang. Study on Suzhou Assessment System of Flood Disaster Condition. Journal of Natural Disasters 1999;14(3):22−26.

[5] Liu Jun, Xu Xiangyang. Flood Disaster Evaluation System of Jiangsu Province. Journal of Catastrophology 1999;8(4):80−83.

The Analysis of Boundary Functions of CMS Reaction Factors

Honghua Xu, Tiejing Li

North East Agriculture University, Harbin, Heilongjiang 150030, China; xhh3161@sohu.com

ABSTRACT: The mathematical model of nonlinear regression was established, describing the effect of sodium hydroxide, monocholroaceetic sodium, and reaction temperature on the synthesis of Carboxymethyl Starch (CMS); and the 23 groups of experiments were carried out. The relationship between the degree of substitution and boundary function of the three factors was discussed. [Nature and Science. 2005;3(2):79-81].

Key words: mathematical model; carboxymethy; boundary function

1 Introduction

Carboxymethyl Starch (CMS), which was first synthesized in 1924, is one of the most important modified starches widely used in textile, petroleum, food and pharmaceutical industries. Different methods for preparing CMS had been reported in literatures [3][4][5]. In the traditional method of synthesizing CMS, the degree of substitution (DS) of CMS is rather low, valued about from 0.01 to 0.1[1]. The products could not be dispersed in cold water, but they are soluble in 80% methanol to permit a reliable determination when the DS is above 0.5[2]. So attention was paid to the reaction ratio to increase DS. In this paper, the relationship between reaction factors of CMS and DS was discussed based on mathematical model.

2 Materials and Methods

2.1 Materials

All chemicals used were of analytical reagent grade; monochloroacetic acid, isopropyl alcohol, alcohol and sodium hydroxide were included; corn starch was produced by the Second Medicine Factory of Harbin.

2.2 Methods

Corn starch (0.5 g) is dissolved in absolute isopropyl alcohol (120 ml), sodium hydroxide was added according to Table 1, and the reaction time is 30 minutes; monochloroacetic sodium (MCAS, $ClCH_2COONa$) was then added (Table 1), and reacted another 60 minutes; then the products were cooled, neutralized, dried, and crushed.

3 Results and Discussion

3.1 Modeling

According to the figures given in the Table 1[6], series of experiments was carried out, using different quantities of aqueous sodium hydroxide and sodium salt of monochloroacetic under different reaction temperatures. The value of DS was got for 23 groups of products (Table 1). So mathematical model of DS of CMS preparations was established, and the three variables are reaction temperature, the dosage of sodium hydroxide and sodium salt of monochloroacetic (Model 1).

$y(x)=0.3451+0.0246x_1+0.1103x_2+0.0428x_3-0.0299 x_1 x_2-0.0138 x_1 x_3+0.0225 x_2 x_3-0.0736 x_1^2-0.0253 x_2^2-0.0258x_3^2$ (Model 1).

F test for Modle1 was carried out, and the result was as followed.

$F_1=1.7010<F_{0.05}(5,8)=3.69$ [6]

$F_2=8.3121>F_{0.01}(9,13)=4.19$

The data of T-test were all higher than $t_{0.4}(1.3)=0.870$ [6]

It was showed that Model 1 was effective, the presence of variances in the equation was of value, and the model reflected the practical situation very well and could well reflect the law of the CMS preparation.

3.2 The relationship between the boundary function of every factors and DS value

Boundary function is that the value of DS changes a Δy with every factor changing a unit of ΔX near a certain level, the contribution of every factors near the level is then showed.

Table 1. Plan of Test and Results

No.	NaOH X_1(g)	Temp. X_2(℃)	MCAS X_3(g)	DS y
1	8	68	14	0.31
2	8	68	6	0.24
3	8	32	14	0.15
4	8	32	6	0.12
5	2	68	14	0.44
6	2	68	6	0.27
7	2	32	14	0.12
8	2	32	6	0.07
9	10	50	10	0.26
10	0	50	10	0.00
11	5	80.3	10	0.49
12	5	19.7	10	0.06
13	5	50	16.8	0.35
14	5	50	3.2	0.19
15	5	50	10	0.35
16	5	50	10	0.36
17	5	50	10	0.28
18	5	50	10	0.35
19	5	50	10	0.23
20	5	50	10	0.39
21	5	50	10	0.37
22	5	50	10	0.33
23	5	50	10	0.34

3.2.1 Analysis of boundary function of NaOH dosage

Reaction temperature (X_2) and $ClCH_2COONa$ dosage (X_3) were fixed in different levels in Model 1, the single factor model of NaOH dosage of were got, and the derivative of the model were calculated, the mathematical model of the boundary function of NaOH dosage to DS was gained.

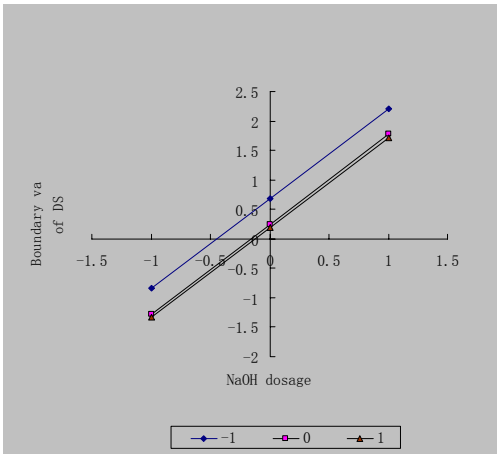


Figure 1. Relationship between DS Boundary Value and NaOH Dosage

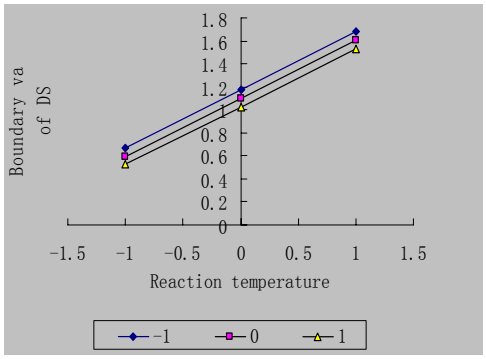


Figure 2. Relationship between DS Boundary Value and Reaction Temperature

When the code value of NaOH dosage was of different level in equation (A), DS of different NaOH dosage was got (Figure 1). The boundary function of NaOH dosage decreased with the increasing of dosage.

The code number of NaOH dosage was of different quantity rank in equation A, and the boundary DS of different NaOH dosages was calculated after number conversion (Figure 1). It was clear that the boundary function of NaOH dosage decreased with increasing of NaOH dosage in Figure 1. The boundary DS was high while NaOH dosage was relatively low; and the boundary DS decreased with the increase of NaOH dosage. The critical point of boundary value of NaOH dosage moved backward with the increasing of other factors, while the NaOH dosage changing from positive value to negative value. When the other factors increased from –1 to 1, their critical point values were around 6.35, 5.48 and 5.36 separately. It was showed that NaOH dosage limited the DS of CMS, and the DS was stable only if every factor was of certain value.

a: $(x_1,-1,-1)$ $dy/dx=-0.1526\ x_1+0.0683$
b: $(x_1, 0, 0)$ $dy/dx=-0.1526\ x_1+0.0246$ (A)
c: $(x_1, 1, 1)$ $dy/dx=-0.1526\ x_1+0.0190$

3.2.2 Analysis of boundary function of reaction temperature

The boundary function of the reaction temperature to the degree of substitution of CMS was in model B and Figure 2.

a: $(-1, x_2,-1)$ $dy/dx=-0.0506\ x_2+0.1177$
b: $(0, x_2, 0)$ $dy/dx=-0.0506\ x_2+0.1103$ (B)
c: $(1, x_2, 1)$ $dy/dx=-0.0506\ x_2+0.1029$

The contributive ratio of reaction temperature to CMS decreased with the increase of temperature, and was positive. It was showed that the increase of reaction temperature do benefit no the synthesis of CMS. When a unit of ΔX_2 was changed in the low temperature

district, the value of DS increased greatly, while it increased slowly in the high temperature district. The degree of DS increasing led by the increase of reaction temperature was weakened when the dosage of $ClCH_2COONa$ and NaOH were increased.

3.2.3 Analysis of boundary function of $ClCH_2COONa$ dosage

The boundary function of the $ClCH_2COONa$ dosage to the degree of substitution of CMS was in model C and Figure 3.

a: $(-1,-1,x_3)$ $dy/dx=-0.0516\ x_3+0.0341$

b: $(0, 0\ x_3)$ $dy/dx=-0.0516\ x_3+0.0426$ (C)

c: $(1, 1,x_3)$ $dy/dx=-0.0516\ x_{31}+0.0515$

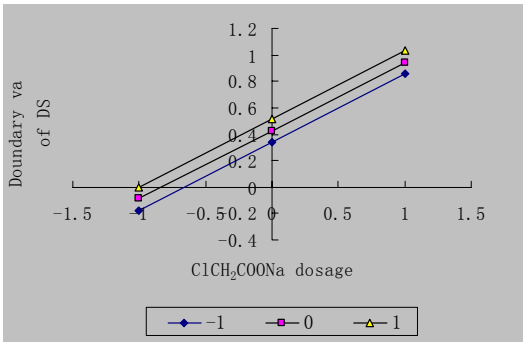


Figure 3. Relationship between DS Boundary Value and $ClCH_2COONa$ Dosage

The change ratio of the degree of subsititution of CMS decreased with the increase of the $ClCH_2COONa$ dosage, and it changed from positive to negetive. That is, the increasing of the $ClCH_2COONa$ dosage did not do benefit on the CMS product with high DS. When other factors were in the level of –1, 0 and 1 separately, the optimum $ClCH_2COONa$ dosage (critical value) was 12.64, 13.32 and 14.00 separately. It was clear that the contributive ratio of the $ClCH_2COONa$ dosage was limited by the taken value of other factors, the critical vale increased with the increasing of the taken value of other factors, and the negative effect led by the increasing of the $ClCH_2COONa$ dosage was decreased.

4 Conclusions

(1) The degree of subsititution of CMS changed with the change of NaOH dosage. It was of the highest value while the NaOH dosage was proper, and lower while the NaOH dosage was too much or small. The limiting function was stable when other factors were fixed.

(2) The reaction temperature was of positive effect to the starch carboxymethylization in the test. The degree of subsititution of CMS increased with the increasing of temperature.

(3) The degree of subsititution of CMS was also affected by the $ClCH_2COONa$ dosage. The $ClCH_2COONa$ dosage had positive effect when the dosage was proper, while it had negative effect when the dosage was too much, but the limiting function decreased with the increasing of other factors' levels.

Acknowledge

Our research is kindly supported by Heilongjiang educational committee.

Correspondence to:
Honghua Xu
Institute of Food Science
Northeast Agriculture University
Harbin, Heilongjiang 150030, China
Telephone: 01186-451-55190479
E-mail: xhh3161@sohu.com

References

[1] Hassan EA, Salama JY. Synthesis and study of polyacids from soluble starch and chloroacetic acids, Starch. 1982;34:375-9.
[2] Khalil MI. Carboxymethyl ethers of corn starch. Starch 1990;42:60-3.
[3] Maronlis ZB. Thermal conductivity of gelatinized starches. J of Food Science 1991;56:773-6.
[4] Moribor VC. Carboxymethl starch. Starch 1972;24:124-6.
[5] Ruenberh MW. Modified dent corn starch. Food Processing 1991;2:102-9.
[6] Xu Zhongru. Regression Analysis and Test Design. China Agriculture Publishing Company, 1997.

Two Probability Theories in Physics

Tianrong Tan

Department of Physics, Qingdao University, Qingdao, Shandong 266071, China, zhangds12@hotmail.com

Abstract: It is proved that in a double slit diffraction experiment, superposition principle of probability amplitude means (the probability amplitude of the event) that an electron passes through a certain slit and arrives somewhere on the screen under the condition that two slits open simultaneously, is equal to the sum of two probability amplitudes of the same event under the condition that two slits open in turn. From this thesis it is concluded that the probabilities do not obey superposition principle, which is just the reason that classical probability theory is inapplicable for the very experiment. Kolmogorov's probability theory is based on two foundations: frequency definition of probabilities and Boolean algebra of event operations. In micro processes, the former holds true while the latter, especially its multiplicative permutation law is violated. So long as we do not deal with event operations, joint probabilities holds good in micro processes. But, the event operation formulae provided are applied and the wrong conclusions present probably. The famous Bell's inequality is a wrong conclusion. [Nature and Science. 2005;3(2):82-91].

Key words: double slit diffraction experiment; Stern-Gerlach experiment; probability amplitudes; Boolean algebra; classical probability theory

Introduction

It is well known that there are two kinds of probabilities in physics. Classical probabilities are applicable for macrophysics while quantum probabilities for microphysics. Hereon, the signification and the scope of application of these two probabilities will be reexamined.

1 Mystery in the Double Slit Diffraction

Famous American physicist R. P. Feynman[1] said that: The double slit diffraction experiments show the whole mystery in quantum mechanics. So far as we know, this mystery can boil down to one conclusion: Classical probability theory is inapplicable to the micro processes. However, it remains an open question that which proposition in classical probability theory is invalid on this occasion and why it is invalid. This question will be examined hereinafter.

To clear the mystery Feynman said, let us examine the following electron double slit diffraction processes.

Firstly, opening the two slits simultaneously, consider that an electron, say e, emitted from the source and has arrived on the screen. Let the E denote that the e passes through the first slit and the F that through the second slit. Then, both E and F are stochastic events. Because that the e has arrived on the screen, it must pass one and only one of the two slits, namely, we have

$$E + F = U \text{ (necessity event)},$$
$$E \cdot F = \emptyset \text{ (impossible event)}. \qquad (1)$$

Secondly, let the sign Ω denote a small area on the screen, and the X denote the event that the e falls on the Ω, hence, $E \cdot X$ ($F \cdot X$) denotes that the e passes through the first (second) slit and falls on the Ω. According to the Boolean algebra of proposition calculation, from Eq, (1) it is concluded that:

$$(E \cdot X) \cdot (F \cdot X) = \emptyset, \qquad E \cdot X + F \cdot X = X.$$

Thirdly, by probability frequency definition, the above formulae give that:

$$\Pr(X) = \Pr(E \cdot X) + \Pr(F \cdot X). \qquad (2)$$

That is a special form of the probability additional formula.

Fourth, by the probability multiplication formula, we have

$$\Pr(E \cdot X) = \Pr(E) \cdot \Pr(X|E);$$
$$\Pr(F \cdot X) = \Pr(F) \cdot \Pr(X|F).$$

Substitute these two formulae into Eq. (2), we have

$$\Pr(X) = \Pr(E) \cdot \Pr(X|E) + \Pr(F) \cdot \Pr(X|F). \quad (3)$$

That is "total probability formula".

To concision, the above proof for Eq. (3) will be written by P-proof in the following. No matter how to examine repeatedly the four steps in this proof, it seems that there is with no chink.

According to the definition, if only open the first slit, the probability that the e fall on the Ω is $Pr(X|E)$; similarly, if only open the second slit, that probability is $Pr(X|F)$. Also, if both slits are open simultaneously, the probability of the very event is $Pr(X)$. According to the total probability formula, $Pr(X)$ is equal to the mean value for $Pr(X|E)$ and $Pr(X|F)$ in accordance with the ratio of $Pr(E)$ and $Pr(F)$. Particularly, if $Pr(E) = Pr(F) = 1/2$, $Pr(X)$ is the arithmetic mean of $Pr(X|E)$ and $Pr(X|F)$. So, the probability that the e falls on the Ω under the condition that the double slit open simultaneously is equal to the arithmetic mean of the two probabilities of the same event under that the double slit open in turn. As a result, the diffraction pattern obtained under the condition that two slits open simultaneously is the overlapping of two diffraction patterns obtained under the condition that two slits open in turn. But the experimental facts have refuted this conclusion: the diffraction patterns under the above two conditions are widely different. It seems certain that this fact indicates: The total probability formula is inapplicable for the double slit diffraction processes.

Now, a contradiction appears. On the one hand, according to the unassailable theory consequence, the total probability formula must be applicable for any a process; on the other hand, the experiment facts indicate the very formula cannot use for the double slit diffraction process. That is just the mystery in quantum mechanics what Feynman said.

2 Double Diffraction and Interpretations for Quantum Mechanics

People are forced to rack their brains for explaining the proposition (a); as a result, many unusual interpretations for quantum mechanics are established.

The first one is the Copenhagen interpretation, of which the main point is as follows: To derive Eq. (3), it needs to affirm Eq. (1). This formula means that the e either passes the first slit or passes the second slit. Copenhagen school demur at this step, the reason is that in the first experiment, we cannot determine which slit the e passed. For Copenhagen school, this fact means that the e neither passes the first slit nor passed the second slit. However, provided e moves along an orbit, it must passes one slit. So, Copenhagen school comes to the conclusion that the electron movement is not orbit movement. And go a step further; they assert that the concepts like "orbit" are "classical concepts" which are inapplicable for micro processes. As such, the Copenhagen interpretation makes a denial of the first step of the P-proof. Because that the first step is unacceptable, the whole proof is invalid naturally.

Quantum logic interpretation (one kind of it) predicates: in proposition calculation the distributive law

$$(E+F)\cdot X = E\cdot X + F\cdot X$$

is inapplicable for this process, and thereby from $E + F = U$ we cannot obtain $E\cdot X + F\cdot X = X$. Hence, this interpretation validates the first step of the P-proof, but negates the second step and so are the after steps.

Both the above two interpretations affirm that the total probability formula and thereby classical probability theory is applicable for micro process, Copenhagen interpretation traces back this promise to classical concept, while quantum logic to classical logic.

Another kind of interpretations only negates classical probability theory itself. For instance, France physicist G. Lochak[2,3] considers that classical probability theory only applicable for "hidden variables", but due to a certain reason, it cannot be used for the mean value of the measurement outcomes. So, Lochak accepts the first and the second step of the P-proof, but refuses the third step.

L. Accardi[4] , who celebrated for establishing quantum probability interpretation, raised a new point: all the preceding three steps of the P-proof hold true so that the addition formula in classical probability theory is applicable for any processes. The trouble is with the fourth step, namely, with the probability multiplication formula, which Accardi calls "Bayes axiom". He said: all of the paradoxes in quantum mechanics result from the improperly using this axiom.

The existence of the above interpretations indicates that: 1. in the P-proof, there is a promise, which is regarded as perfectly justified but is applicable for micro process really. 2. it is still an open question what such a

promise is.

3 A Hidden Promise

As we see, to explain the (a) various interpretations for quantum mechanics has been advanced. Each of them abandoned a specific promise of P-proof. Concretely speaking, four promises are given up respectively: the classical concept about orbit; the contribution law in proposition calculation; the probability addition formula and the probability multiplication formula.

English philosopher Karl Popper put forward a new point about double slit diffraction experiment as follows[5]: Every change of the experiment devices, for instance, closing a slit, will make an impact on the distribution of various possibilities. This point made the pivotal step for revealing the above mystery.

In the P-proof, all the four promises above mentioned are necessary, but Popper opened out another promise that is overlooked all the way: In the proving Eq. (2), the $Pr(E\cdot X)$ and $Pr(F\cdot X)$ are the probabilities under the condition that two slits are open simultaneously, but when Eq. (3) is used for double diffraction process, they are regarded as the probabilities under the condition that two slits are open in turn. Therefore, when people obtain the (a), it is tacitly approved that the $Pr(E\cdot X)$ (as well as the $Pr(F\cdot X)$) takes the same value under the above two conditions. However, this promise is not an unalterable principle. We must regard it as the fifth promise of the P-proof.

To formulize this promise, let the sign *C* denote the condition that two slits are open simultaneously, and the *D* that open in turn. Then, according to the preceding three steps of the proof, we can only obtain

$$\Pr(X|C) = \Pr(E\cdot X|C) + \Pr(F\cdot X|C), \qquad (4)$$

but the addition formula inapplicable for double slit diffraction process is

$$\Pr(X|C) = \Pr(E\cdot X|D) + \Pr(F\cdot X|D). \qquad (5)$$

Now, let us distinguish Eq. (5) from Eq. (4).

Eq. (4) can be interpreted as follows: Consider a process, two slits are opened for a time interval T, there are N electrons falling on the screen and n one falling on the Ω. Among those n electrons, n_1 one passed through the first slit and n_2 one passed through the second slit. Due to the number N is large enough, the probability in the left side of Eq. (4) is n/N; and those in the right side of Eq. (4) are respectively n_1/N and n_2/N, so that Eq. (4) indicates

$n = n_1 + n_2$, which is a self-evident outcome.

Also, Eq. (5) can be interpreted as follows: Consider another process, under otherwise identical conditions, at the beginning only the first slit is opened for a time interval T, there are m_1 electrons falling on the Ω; and later on only the second slit is opened for an equal time interval and m_2 electrons falls on the Ω. Consequently, the probabilities in the right side of Eq. (5) are m_1/N and m_2/N respectively, so that Eq. (5) indicates: $n = m_1 + m_2$, *(6),* which is just the fifth promise of P-proof.

Now, let us explain the meaning of this promise, when two slits open for the time interval T, the electrons passing through the first slit form an electron beam, say C_1; and another electron beam, say C_2, is formed by the electrons passing through the second slit. In the first process above mentioned, the C_1 and the C_2 arrive the screen together, while in the second process, the C_1 and the C_2 arrive the screen one after the other. As a result, the fifth promise can be expressed as follows: The electron number falling on the Ω under the condition that the C_1 and the C_2 arrive the screen simultaneously, is equal to two electron numbers falling on the Ω under the condition that the two beams arrive the screen in turn. Expressed by the probabilistic form, it becomes that: In the double slit experiment, the probability of the event that an electron falls on the Ω under the condition that two slits open simultaneously, is equal to two probabilities of the same event under the condition that two slits open in turn.

4 Superposition Principle

According to electrostatics, if there are two point charges, and field intensity at observe point is $\boldsymbol{E}_1$ if only the first change exists, and is $\boldsymbol{E}_2$ if only the second change exists, then that is $\boldsymbol{E}_1 + \boldsymbol{E}_2$ if both two charges exist. This fact is called after superposition principle about electrostatic field. From this instance, the general definition for superposition principle can be described as follows:

Definition 1: Assume that the presence of something will produce a certain effect. If in a process, the effect produced by the presence of two such things is equal to the sum of the effects of the separated presence

of each thing, then we say in the very process, the effects of the things obey superposition principle.

Compared with this definition, it is seen that the proposition (b) indicates that the probabilities obey superposition principle. But, because that the (b) is not an experimental fact, it cannot be named as a principle, so that we call it after "probability superposition assumption" herein. Hence, the conclusion we obtained from the double slit experiment is that the probability superposition assumption is invalid.

To manifest this conclusion in a more universal form, two terms will be introduced as follows: Firstly, the conditions such as the C and the D in called by "Popper conditions". Secondly, following quantum mechanics, the events like "the e passing through the first slit falls on the Ω" are called as "transition events". As such, the above conclusion is generalized as follows:

In a transition process, for Popper conditions, the probabilities do not obey superposition principle.

Adopting the above symbols, Eq. (6) is equivalent to:

$$n_1 = m_1, \qquad n_2 = m_2. \qquad (7)$$

Rewriting the $Pr(A|C)$ as $Pr_C(A)$ and the $Pr(A|D)$ as $Pr_D(A)$, as a result of the probability frequency definition, the total probability formula is

$$Pr_C(X) = Pr(E) \cdot Pr_C(X|E) + Pr(F) \cdot Pr_C(X|F), \quad (8)$$

but the formula negated by facts is

$$Pr_C(X) = Pr(E) \cdot Pr_D(X|E) + Pr(F) \cdot Pr_D(X|F), \quad (9)$$

which is a formula possessing the classical probability theory characters. The transition from Eq. (8) to Eq. (9) requires the following relations:

$$Pr_C(X|E) = Pr_D(X|E), \quad Pr_C(X|F) = Pr_D(X|F). \quad (10)$$

These equations are just that is the probability expression for Eq. (7), which signify the fifth promise.

It seems that all of the preceding four promises are unassailable; so as to no matter giving up any one of them, we will result in certainly some inconceivable outcomes. But the fifth promise does not so; it is far not perfectly justified. Merely due to carelessness, it is accepted. As viewed from the other angle, the discovery of this promise is by means of careful analysis instead of by a certain bold new idea.

Eq. (10) can be expressed as follows: Under the known condition that the e passes through a certain slit, the probability that it falls on the Ω is independent of the condition whether or not the other slit is open. Hereon, the condition such as the other slit is open is another form of Popper condition. So that, the above proposition can be generalized as that: Under the change of Popper conditions, the transition probabilities hold constant. This is another form for the fifth promise in P-proof. Give up this promise, classical concept such as orbit movement; classical logic such as the distributive law of proposition calculation; classical probability theory laws such as addition formula and multiplication formula, all can be saved from abandon.

The (c) is applicable to various processes; especially, it is applicable to the Stern-Gerlach experiment that we will examine in the following.

5 Stern-Gerlach Experiment

When an electron beam passes through a non-uniform magnetic field with orientation $\boldsymbol{n}$, it will split apart into two sub-beams. In one of them, which we call the beam magnetized along the $\boldsymbol{n}$, the projections in the $\boldsymbol{n}$ direction of the electron spins will be $\sigma_n = 1$ (measured by h/4π); in the other, we call the beam magnetized in the $-\boldsymbol{n}$, $\sigma_n = -1$.

Assume that there is a beam having N electrons magnetized along the $\boldsymbol{n}$, after passing through a magnetic field orientated by $\boldsymbol{m}$, there are M electrons magnetized along the $\boldsymbol{m}$, then we say the probability that a single electron in the very beam transits from the state $\sigma_n = 1$ to the $\sigma_m = 1$ is M/N (if the N is large enough), which is written as a conditional probability $Pr(\sigma_m = 1 \mid \sigma_n = 1)$, and is called a "transition probability".

Feynman described a kind of perfect Stern-Gerlach devices (abbreviated by "devices" below), which possess the following properties: One of such devices has a character direction, after entering into it a beam will split to two sub-beams, the one is magnetized along the character direction and the other along the reverse direction. Besides, two devices can be linked up such that all electrons escaping from the preceding device can enter into the behind one. In each device, there is a baffle that may absorb one of the two sub-beams. In the following it is stipulated that a device is written as G_n if its character direction is $\boldsymbol{n}$.

To compare the Stern-Gerlach experiment with the

double slit diffraction experiment, we consider a special process, which we call "character process", as follows: Assume that a beam, say R, passes through a device G_a and splits apart into two sub-beam, the one is absorbed in it and the other, say A, escapes from it. Then, the A, which has N electrons and is magnetized along the $\boldsymbol{a}$, enters into another device G_c, and splits apart into two sub-beam C_1 and C_2. The C_1 has N_1 electrons and is magnetized along the $\boldsymbol{c}$ and the C_2, N_2 electrons and magnetized along the $-\boldsymbol{c}$. Afterwards, C_1 and C_2 depart from the G_c and enter into the third device G_b together. In the G_b, the entered electrons realign and become two sub-beams, one of them, say B, magnetized along $\boldsymbol{b}$, leave the G_b and the other is absorbed in it.

In this process, after entering into the G_c, an electron may belong to C_1 as well as to C_2, which are two passages through the G_c. We call the preceding one after the first passage and the other the second passage. Assume that in the A, there are M_1 electrons passing through the first passage of the G_c and finally departing from G_b; and M_2 electrons through the second passage and departing from the G_b, and thereby there are $M_1 + M_2$ electrons in the A finally belonging in the B.

Let e denote an electron in the A, then, according to the frequency definition, under the condition that the N is large enough, the probability that the e passes the first passage is $\Pr(\sigma_c = 1 \mid \sigma_a = 1) = N_1 / N$; when the condition that the e passes the first passage is known, the probability that it departs from G_b is

$$\Pr(\sigma_b = 1 \mid \sigma_c = 1) = M_1 / N_1,$$

So, the probability that the e passes through the first passage of the G_c and finally departs from the G_b is

$$p_1 = \Pr(\sigma_c = 1 \mid \sigma_a = 1) \cdot \Pr(\sigma_b = 1 \mid \sigma_c = 1) = M_1 / N. \quad (11)$$

Similarly, the probability that the e through the second passage of the G_c and finally departs from the G_b is

$$p_2 = \Pr(\sigma_c = -1 \mid \sigma_a = 1) \cdot \Pr(\sigma_b = 1 \mid \sigma_c = -1) = M_2 / N. \quad (12)$$

According to definition, the probability that the e finally departs from the G_b is: $p = (M_1 + M_2) / N$.

Eq. (11), Eq. (12) and the above formula give

$$p = p_1 + p_2. \quad (13)$$

On the other hand, according to the experiment facts, if the e enters into the G_b directly (namely, it does not pass through the G_c), the probability that it departs from the G_b is also the p. So, $p = \Pr(\sigma_b = 1 \mid \sigma_a = 1)$. (14)

Substitute Eq. (11), Eq. (12) and Eq. (14) into Eq. (13), we obtain that:

$$\Pr(\sigma_b = 1 \mid \sigma_a = 1) = \sum_z \Pr(\sigma_c = z \mid \sigma_a = 1) \cdot \Pr(\sigma_b = 1 \mid \sigma_c = z),$$

where $z \in \{1, -1\}$.

Generally speaking, for any $x, y, z \in \{1, -1\}$,

$$\Pr(\sigma_b = y \mid \sigma_a = x) = \sum_z \Pr(\sigma_b = y \mid \sigma_c = z) \cdot \Pr(\sigma_c = z \mid \sigma_a = x). \quad (15)$$

It should be noted that this formula is only significant for the character process. Generally speaking, in micro world, a formula consisting of probability expressions are only significant for a given process, because that probability superposition assumption is valid.

6 Comparison Between two Total Probability Formulae

To compare the Stern-Gerlach experiments with the double slit diffraction experiments, let S denote the condition that the e is falls on the screen, this is the precondition condition for the later process, hence for this process the total probability formula

$$\Pr(X) = \Pr(E) \cdot \Pr(X|E) + \Pr(F) \cdot \Pr(X|F),$$

is rewritten as

$$\Pr(X|S) = \Pr(E|S) \cdot \Pr(X|E) + \Pr(F|S) \cdot \Pr(X|F).$$

Herein, the S is the collection of the electrons, which pass through one of the double slit. This is a beam corresponding with the beam A in the Stern-Gerlach experiment, namely, the beam $\sigma_a = 1$. Also, the E also denotes a beam, of which the electrons pass through the first slit. It is corresponding with the beam C_1 in the Stern-Gerlach experiment, and the X, a beam with electrons fall on the Ω, is corresponding with the beam B. It is thus seen that Eq. (15) is corresponding to the total probability formula for the transition probability $Pr(\sigma_m = 1 \mid \sigma_n = 1)$.

According to definition, all of p_1, p_2 and p in Eq. (11) are probabilities under the condition that two passages open simultaneously. The p is the ratio of the number of the electrons departing from the G_b to that entering into the G_c, which is measurable; but the p_1 and the p_2 is not. To measure them it is need to consider another experiment. Provided that after the A entering into G_c and splitting to C_1 and C_2, absorbing the C_2 by

the baffle in the G_c, we can measure the p_1. Similarly, we can also measure the p_2. However, the p_1 and p_2 obtained in such a way are the probabilities under the condition that two slits open in turn. The values of them may be different from those in Eq. (11) and Eq. (12), because those are obtained from the distinct Popper conditions.

Such being the case, in Eq. (15), which probabilities are dependent on Popper conditions? So-called Popper conditions herein means under otherwise identical conditions let the G_c open two passages simultaneously or in turn. So, the p in Eq. (14) is independent of Popper conditions. The p_1 given by Eq. (11) has two factors, in them the $Pr(\sigma_c = 1 \mid \sigma_a = 1)$ is Popper condition free because that the splitting of the A in the G_c is before than the C_1 or C_2 is absorbed. The unique probability herein dependent on Popper conditions is the factor $Pr(\sigma_b = 1 \mid \sigma_c = 1)$. Similarly, all the probabilities in the form $Pr(\sigma_b = y \mid \sigma_c = z)$ depend on Popper conditions. Therefore, such expressions have twofold meaning. In the following, we will distinguish them by symbols.

Using the symbols Pr_C and Pr_D to distinguish the probabilities under conditions two passages open simultaneously and that in turn, as a formula obtained by frequency definition, Eq. (15) ought to be expressed as that:

$$\Pr(\sigma_b = y \mid \sigma_a = x) = \sum_z \Pr_C(\sigma_b = y \mid \sigma_c = z) \cdot \Pr(\sigma_c = z \mid \sigma_a = x). \quad (15a)$$

but if the p_1 and p_2 is regarded as the measurement value, it becomes

$$\Pr(\sigma_b = y \mid \sigma_a = x) = \sum_z \Pr_D(\sigma_b = y \mid \sigma_c = z) \cdot \Pr(\sigma_c = z \mid \sigma_a = x). \quad (15b)$$

This formula manifests the superposition assumption of probabilities in the very process; it is a relation between the measurement values. However, this relation is invalid.

7 Probability Amplitudes

As we see, by means of the probability frequency definition we obtain Eq. (15); but, because that the probability superposition assumption is invalid, the frequency definition is incompetent for finding the relation between the measurable probabilities in Stern-Gerlach experiment. To this end, we must find a quantity in micro processes, which obey superposition principle and is able to calculate the probabilities. Fortunately, such a quantity has been found, it is called "probability amplitude".

Feynman said that the concept of probability amplitude is the core of quantum mechanics. Actually, the importance of probability amplitude just rests with that it obeys the superposition principle about the Popper conditions. For the double slit diffraction experiment or the like of it, the superposition principle for probability amplitude can be expressed as follows:

Assume that there are two passages to transit from A state to B state, the probability amplitude of the event that a single electron from the A arrives at the B when two passages open simultaneously is equal to the sum of two probability amplitudes of the same event when two passages open in turn.

Also, quantum mechanics gives that:

The corresponding relation between the transition probability $Pr(B \mid A)$ and its amplitude, which is written as $\langle B \mid A \rangle$, is $\Pr(B \mid A) = |\langle B \mid A \rangle|^2$.

Similar to probabilities, the amplitudes also satisfy addition formula and multiplication formula. By means of addition formula, superposition principle can be rewritten as follows: Under different Popper conditions, the probability amplitudes for a given transition event is the same.

According to quantum mechanics, for arbitrary given unit vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\gamma = \angle(\boldsymbol{a}, \boldsymbol{b})$, the probability amplitude $\langle \sigma_b = y \mid \sigma_a = x \rangle$ takes the values as follow:

$$\langle \sigma_b = 1 \mid \sigma_a = 1 \rangle = \langle \sigma_b = -1 \mid \sigma_a = -1 \rangle = \cos\frac{\gamma}{2};$$

$$\langle \sigma_b = -1 \mid \sigma_a = 1 \rangle = \langle \sigma_b = 1 \mid \sigma_a = -1 \rangle = i\sin\frac{\gamma}{2}. \quad (16)$$

Now, Eq. (16) is applied to rewriting Eq. (15b). Consider the process that the beam A entering into the G_b directly. From Eq. (16) and the (B), it is concluded that:

$$\Pr(\sigma_b = 1 \mid \sigma_a = 1) = \Pr(\sigma_b = -1 \mid \sigma_a = -1) = \cos^2\frac{\gamma}{2},$$

$$\Pr(\sigma_b = 1 \mid \sigma_a = -1) = \Pr(\sigma_b = -1 \mid \sigma_a = 1) = \sin^2\frac{\gamma}{2}. \quad (17)$$

Next, think about the process that C_1 and C_2 enter into the G_b in turn. To simplify the question, we assume

that ***a***, ***b***, ***c*** are in one plan and $\angle(\boldsymbol{a}, \boldsymbol{b}) = \angle(\boldsymbol{b}, \boldsymbol{c}) + \angle(\boldsymbol{a}, \boldsymbol{c})$.

Let

$\angle(\mathbf{b}, \mathbf{c}) = \alpha$, $\angle(\mathbf{a}, \mathbf{c}) = \beta$, $\angle(\mathbf{a}, \mathbf{b}) = \gamma = \alpha + \beta$.

By Eq. (17), the probability of the event that the e, a single electron in A, obtain the measurement value $\sigma_c = 1$ in G_c is $\cos^2(\beta/2)$, and when this condition has been known, the probability that the e obtain the measurement value $\sigma_b = 1$ in G_b is $\cos^2(\alpha/2)$. As a result, when only the first passage open, the e passes the G_c and finally obtains measurement value $\sigma_b = 1$ is

$$q_1 = \cos^2\frac{\alpha}{2}\cos^2\frac{\beta}{2}.$$

Similarly, when only the second passage open, it passes the G_c and finally obtains $\sigma_b = 1$ is

$$q_2 = \sin^2\frac{\alpha}{2}\sin^2\frac{\beta}{2}.$$

When C_1 and C_2 enter into the G_b simultaneously, the probability superposition assumption gives the probability that the e passes through the G_c and the G_b successively and finally obtains $\sigma_b = 1$ is $p = q_1 + q_2$. *(18)*

However, since that the probability superposition assumption is invalid, Eq. (18) is wrong. Fortunately, we can apply Eq. (16) to calculate the p in the very process. As such, the amplitude that the e obtains $\sigma_c = 1$ in the G_c is $\cos(\beta/2)$; also, under the above known condition the amplitude that it obtain the measurement value $\sigma_b = 1$ is $\cos(\alpha/2)$. By the (C), the amplitude that the e along the first passage and finally obtains $\sigma_b = 1$ is $\cos(\alpha/2)\cos(\beta/2)$. Similarly, along the second passage the amplitude of the same event is $-\sin(\alpha/2)\sin(\beta/2)$. Applying the (C) once again, we obtain the amplitude corresponding to p as follows: $\cos\frac{\alpha}{2}\cos\frac{\beta}{2} - \sin\frac{\alpha}{2}\sin\frac{\beta}{2} = \cos\frac{\alpha+\beta}{2}$.

Now, the (B) gives $p = \cos^2\frac{\alpha+\beta}{2}$.

The above formulae give

$$p = q_1 + q_2 - \frac{1}{2}\sin\alpha\sin\beta.$$

Comparing with Eq. (18), it is seen that the $-\frac{1}{2}\sin\alpha\sin\beta$ in the above equality is the "intersection term" that make the probability superposition assumption invalid. Let the sign J denote this intersection term, Eq. (15b) is rewritten as

$$\Pr(\sigma_b = y \mid \sigma_a = x) = \sum_z \Pr_D(\sigma_b = y \mid \sigma_c = z) \cdot \Pr(\sigma_c = z \mid \sigma_a = x) + J.$$

In the other field, the energy density equation of the electrostatic field for two point charges is

$u = u_1 + u_2 + u_{12}$.

In which the u_{12} is the "intersection term" that makes that the energy density of the electrostatic field did not obey the superposition principle. It seems that such comparability indicates that between micro process and macro process there is no impassable chasm.

As to p_1 and p_2, they are immeasurable and thereby cannot be obtained by the calculation of amplitudes.

8 Boolean Algebra and Event Operations

It is well known that the probability theory is founded by A. N. Kolmogorov. In the beginning of this theory, in which there is no superposition assumption of probabilities. On the other hand, when people say that the classical probability theory is inapplicable for micro process, this assumption is regarded as a component of it actually. Considering this practice, we stipulate that the Kolmogorov probability theory is named for the system containing the whole probability operation laws, while the classical probability theory for the conjunction of Kolmogorov probability theory and superposition assumption of probabilities.

It has been seen that due to the existence of the probability superposition assumption probabilities the classical probability theory is not always applicable for micro process. Now, a question appears naturally: whether or not the Kolmogorov probability theory applicable for micro processes completely?

If examining closely, we can see that the Kolmogorov probability theory is based on two foundations: one is frequency definition for probabilities (to derive the multiplication formula from probability frequency definition requires a slight revision for such definition); the other is Boolean algebra for event operations. Frequency definition, which is connatural for probabilities, is universally accepted that it is applicable for micro processes. The question is whether Boolean algebra for event operations applicable for micro processes.

Event operations and proposition calculations are

applications of Boolean algebra in different domains. As we see, it is unacceptable to modify the Boolean algebra as the laws of proposition calculations. However, from this reason it cannot be concluded that Boolean algebra as the laws of event calculations is also unchangeable.

In Kolmogorov probability theory, if A and B are two events, the occurrence of the product event $A \cdot B$ indicates that "the A occurs and the B occurs". So, the occurrence of the $B \cdot A$ indicates that "the B occurs and the A occurs". Though the above two product events are different in expression form, the meanings are the same actually. As a result, the commutable law

$$A \cdot B = B \cdot A \qquad (19)$$

holds true. In other words, in the expression $A \cdot B$ the positions of the A and the B are commutable.

However, if we prescribe that the occurrence of the $A \cdot B$ means that the event "the A occurs before and the B occurs behind", then the occurrence of the $B \cdot A$ means that the event "the B occurs before and the A occurs behind". Such, when commuting the positions of the A and the B in the expression $A \cdot B$, not only the expression form but also the real meaning is changed, and Eq. (19) is thus invalid.

In the micro processes, this situation has been fallen across. For example, in the double slit diffraction process, let the sign E denote that "the e passes the first slit" and the X that "the e falls on the small area Ω on the screen", then the product event $E \cdot X$ denotes that "the e passes the first slit and then falls on the Ω". In fact, the $E \cdot X$ posses such content, so that the multiplication operation herein does not obey the commutable law.

It is thus seen that we must be especially careful in the matter relating to micro processes. For instance, according to frequency definition, the multiplication formula

$$\Pr(A \cdot B) = \Pr(A) \cdot \Pr(B \mid A) \qquad (20)$$

as a Kolmogorov probability theory law is valid for micro processes.

Because that Eq. (20) is an identity, which holds true for the commuting of the independent variables, so that from Eq. (20) we can obtain

$$\Pr(B \cdot A) = \Pr(B) \cdot \Pr(A \mid B),$$

so far as the events in it are meaningful. But due to the commutable law is invalid, from Eq. (20) we cannot obtain $\Pr(A \cdot B) = \Pr(B) \cdot \Pr(A \mid B)$.

As a result, if in the Stern-Gerlach experiment we have, by means of multiplication formula, gotten

$$\Pr(\sigma_a = x,\ \sigma_b = y) = \Pr(\sigma_a = x) \cdot \Pr(\sigma_b = y \mid \sigma_a = x), \qquad (21)$$

then, because that the commutable formula

$$\Pr(\sigma_a = x,\ \sigma_b = y) = \Pr(\sigma_b = y,\ \sigma_a = x)$$

is invalid generally, from Eq. (21) it is impossible to obtain

$$\Pr(\sigma_b = y,\ \sigma_a = x) = \Pr(\sigma_a = x) \cdot \Pr(\sigma_b = y \mid \sigma_a = x).$$

In Boolean algebra, some more complex formula, for example, $(A \cdot B) \cdot C = (A \cdot C) \cdot B$, is invalid in the micro processes because that the commutable law has been used.

9 Joint Probabilities

A probability for product of several events is called after "joint probability". It is well known that the very concept is useless in quantum mechanics, but this fact does not mean that the application of joint probabilities herein will certainly go wrong. According to frequency definition, it is possible to introduce joint probabilities without conflicts. But we must keep in mind that the joint probabilities herein may not obey the laws of Boolean algebra for event operation. Otherwise, the mistakes will appear. As an example, we consider a formula playing an important part in so-called Bell's theorem.

Twice applying the multiplication formula, we have

$$\Pr(A \cdot B \cdot C) = \Pr(A) \cdot \Pr(B \mid A) \cdot \Pr(C \mid A \cdot B). \qquad (22)$$

If the $A \cdot B \cdot C$ is regarded as the product of three events occurring successively, this formula is applicable for micro processes. On the other hand, we know that the electrons "lack of memory", which means that in the character process the expression $Pr(\sigma_b = y \mid \sigma_c = z, \sigma_a = x)$ can be abbreviated to $Pr(\sigma_b = y \mid \sigma_c = z)$. Therefore, for joint probability $Pr(\sigma_a = x, \sigma_b = y, \sigma_c = z)$, Eq. (22) gives

$$\Pr(\sigma_a = x,\ \sigma_b = y,\ \sigma_c = z) = \Pr(\sigma_a = x) \cdot \Pr(\sigma_c = z \mid \sigma_a = x) \cdot \Pr(\sigma_b = y \mid \sigma_c = z). \qquad (23)$$

By means of Eq. (23), Eq. (15) becomes

$$\Pr(\sigma_a = x,\ \sigma_b = y) = \sum_z \Pr(\sigma_a = x,\ \sigma_b = y,\ \sigma_c = z)$$

.(24)

This formula can be understood by the addition formula for Stern-Gerlach process.

In Eq. (24) commuting the $\sigma_b = y$ and $\sigma_c = z$, we have

$$\Pr(\sigma_a = x, \sigma_c = z) = \sum_y \Pr(\sigma_a = x, \sigma_c = y, \sigma_b = z)$$

.(25)

The Kolmogorov probability theory gives

$$\Pr(\sigma_a = x, \sigma_b = y, \sigma_c = z) = \Pr(\sigma_a = x, \sigma_c = y, \sigma_b = z)$$

(26)

by which Eq. (25) becomes

$$\Pr(\sigma_a = x, \sigma_c = z) = \sum_y \Pr(\sigma_a = x, \sigma_b = y, \sigma_c = z).$$

In a similar way, we can obtain

$$\Pr(\sigma_b = y, \sigma_c = z) = \sum_x \Pr(\sigma_a = x, \sigma_b = y, \sigma_c = z).$$

Writing $Pr(\sigma_a = x, \sigma_b = y, \sigma_c = z)$ as $F(x, y, z)$, from Eq. (24) and the above two formulae, it is concluded that:

For any given directions $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and $x, y, z \in \{1, -1\}$, there exists function $F(x, y, z) \geqslant 0$, such that:

$$\Pr(\sigma_a = x, \sigma_b = y) = \sum_z F(x, y, z);$$

$$\Pr(\sigma_a = x, \sigma_c = z) = \sum_y F(x, y, z);$$

$$\Pr(\sigma_b = y, \sigma_c = z) = \sum_x F(x, y, z).$$

This is a proposition resulting from Kolmogorov probability theory, is it valid? Let us check it step by step.

Firstly, in the proof for the (d), Eq. (22) has been used. By frequency definition we know that Eq. (22) is hold true for the character process.

Secondly, by multiplication formula, from Eq. (22) we get Eq. (24). Due to multiplication formula is valid in micro process, the Eq. (24) is also hold true.

Thirdly, in the both sides of Eq. (24), commuting $\sigma_b = y$ and $\sigma_c = z$, Eq. (25) is obtained, this step is reasonable for identity, so that Eq. (25) is also valid.

Fourth, the deducing the (d) from Eq. (25), the Kolmogorov probability theory formula Eq. (26) is used, this step is unreasonable however. The left side of Eq. (26) is the probability that in a character process that the e passes through the G_c before and through the G_b behind, while the right side of Eq. (26) is the probability that under otherwise equal conditions, commuting G_c and G_b, namely, the e passes through the G_b before and through the G_c behind. Clearly, the latter is different from the farmer naturally. So, Eq. (26) is certainly invalid.

It is thus concluded that the proposition (d) is wrong.

It should be noted that: Firstly, Eq. (26) is independent of Popper conditions, so that the invalidity of Eq. (26) has nothing to do with probability superposition assumption. Secondly, the Kolmogorov probability theory is based on the frequency definition of probabilities and Boolean algebra of event operations. Eq. (26) is invalid because of Boolean algebra of event operations instead of frequency definition of probabilities.

10 Micro Processes and Probability Operations

Because that probability amplitudes obey superposition principle about Popper conditions, provided that apply probability amplitudes to calculating probabilities, it is naturally to be concluded that the probabilities do not obey superposition principle about Popper conditions. Besides, probability amplitude involves two states, one is the state before the transition and the other is behind the transition. These two state are unsymmetrical, it is also naturally to be concluded that the multiplication operation of events do not obey the commutable law. As a result, the application of probability amplitude naturally removes the two factors those are inapplicable for micro processes. So, even if one is ignorant of the character of the probability operations in micro processes, provided holding the techniques of operating the probability amplitude, he is able to gallop across the micro field. This instance is especially lucky for quantum physicists. But there is a small deficiency: when the problem not only involves the techniques but also relates to the substance of probability calculations, they are hard to avoid suffering setback. The muddle brought about by Bell's inequality, which will be examined in another paper, is a fact illustrative of the very point.

Correspondence to:
Tianrong Tan
Department of Physics
Qingdao University
Qingdao, Shandong 266071, China
zhangds12@hotmail.com

References

[1] Feynman RP. The Feynman Lectures on Physics, Vol. 3.

[2] Lochak G. Has Bell's Inequality a General Meaning for Hidden-Variable Theories? [J]. Foundations of Physics, 1976;6 (9).

[3] Lochak G. De Broglie's Initial Conception of de Broglie Waves [A]. Diner S et al. (ads.) The Wave - Particle Dualism [M]. D. Reidel Publishing Company, 1984.

[4] Accardi L. The Probabilistic roots of the quantum mechanical paradoxes [A]. S. Diner et al.(ads.), The Wave - Particle Dualism[M]. 1984:297 - 330.

[5] Popper KR. The propensity interpretation of the calculus of probability, and quantum mechanics, [M] Observation and Interpretation in the Philosophy of Physics S. Korner & M. H. L. Pryce (pp.65-70).

A Concise Introduction of Avian Influenza and Its Pathogen: the Requirement of Public Health

Guangxing Li [1], Jiechao Yin [2], Binjie Wang[3],Yudong Ren [4], Xiaofeng Ren [2]

1. Department of Basic Veterinary, College of Veterinary Medicine, Northeast Agricultural University, 59 Mucai Street, Harbin, Heilongjiang 150030, China; Email: ligx@mail.neau.edu.cn
2. Department of Preventive Veterinary, College of Veterinary Medicine, Northeast Agricultural University, 59 Mucai Street, Harbin, Heilongjiang 150030, China; Telephone: 0086-451-55190385; Email: rxfemail@yahoo.com.cn, crane198026@hotmail.com
3. The Clinic of Ophthalmology, The 1st Affiliated Hospital of Harbin Medical University, 143 Yiman Street, Harbin, Heilongjiang 150001, China
4. Department of Computer, College of Engineering, Northeast Agricultural University, 59 Mucai Street, Harbin, Heilongjiang 150030, China; Email: rxfemail@yahoo.com.cn

Abstract: In the late 1990s, types of H5N1 and H9N2 avian influenza viruses (AIV) have infected respiratory tract in humans in Hong Kong through the domestic poultry to the potential contact with human being, which highlights the status of public health of veterinary pathogen. Particularly, the recent outbreak of avian influenza (AI) together with the temporary SARS prevalence and the natural disaster caused by tsunami in Asia, impacting on the economic and society significantly. Therefore, we emphasize here the importance of AIV and depict the etiology, molecular biology, clinical signs, diagnosis and prevention of AI in order to update the scientific literatures, increase the understanding to AIV, improve the prevention consciousness in public. [Nature and Science. 2005;3(2):92-94].

Key words: avian influenza; public health; pathogen; prevention

1 Why Avian Influenza Viruses Become the Spotlight in Public?

Influenza A viruses own 15 hemagglutinin (HA) and 9 neuraminidase (NA) subtypes, existing in waterfowl extensively, is the source of influenza A viruses that infect mammals. Historically, only influenza A H1N1, H2N2 and H3N2 viruses have caused the pandemics in Human (1918, 1957, and 1968 respectively). Normally, pandemic strains possessing novel HA derived from avian or animal influenza viruses, regardless of other accompanying avian virus genes, sporadically emerge in humans and have the potential to cause a pandemic of influenza if the virus is capable of transmitting among a human population that lacks immunity to the novel HA [1]. Generally, avian influenza viruses were not considered to be able to directly infect humans and cause influenza-like respiratory disease due to the discrepancy of specific receptor and a potential barrier within species. However, A high-pathogenicity avian H5N1 influenza viruses circulated among poultry in farms and retail markets in 1997, in Hong Kong (The H5N1 viruses were subsequently transmitted to humans, causing 18 documented cases of respiratory disease, including 6 deaths). Likewise, in 1998–99, a second influenza A virus subtype, H9N2, was isolated from humans with respiratory disease. More recently, some lethal cases have been reported in Asian [2]. Now, there are less doubt that these avian influenza viruses should be considered a risk to public health and social economy.

2 Etiology

The AIV belong to the virus family rthomyxoviridae. The virus particle has an envelope with glycoprotein projections with hemagglutinating and neuraminidase activity. These two surface antigens, HA and NA, are the serologic taxonomy basis of the influenza viruses using the letters H and N with the appropriate numbers in the virus designation e.g., H7N2.

Currently there are 15 hemagglutinin and 9 neuraminidase antigens described among the Type A influenza viruses. The type designation (A, B, or C) is based upon the antigenic character of the M protein of the virus envelope and the nucleoprotein within the virus particle. All influenza viruses affecting domestic animals (equine, swine, avian) belong to Type A, and Type A influenza virus is the most common type producing serious epidemics in humans. Types B and C do not affect domestic animals [3].

3 Molecular Basis of Pathogenicity

Influenza A viruses the hemagglutinin glycoprotein is produced as a precursor, HA0, which requires posttranslational cleavage by host proteases before it is functional and virus particles are infectious. The HA0 precursor proteins of low-pathogenicity avian influenza (LPAI) have a single arginine at the cleavage site and another basic amino acid at position -3 or -4. These viruses are limited to cleavage by extracellular host proteases such as trypsin-like enzymes and thus restricted to replication at sites in the host where such enzymes are found, i.e., the respiratory and intestinal tracts. High-pathogenicity avian influenza (HPAI) viruses possess multiple basic amino acids (arginine and lysine) at their HA0 cleavage sites either as a result of apparent insertion or substitution and appear to be cleavable by a specific protease(s), probably one or more proprotein-processing subtilisin-related endoproteases of which furin is the primary candidate. HPAI viruses can replicate throughout the bird, damaging vital organs and tissues, which result in disease and death [4].

4 Epidemiology

Most of all, AI outbreaks probably start with direct or indirect contact of domestic poultry with waterbirds. Many strains that circulate in wild birds are either non-pathogenic or mildly pathogenic for poultry. However, a virulent strain may emerge either by genetic mutation or by reassortment of less virulent strains. Swine appear to be important in the epidemiology of infection of turkeys with swine influenza virus when they are in close proximity. Other mammals do not appear to be involved in the epidemiology of HPAI. The infection of humans with an H5 avian influenza virus in Hong Kong in 1997 has resulted in a reconsideration of the role of the avian species in the epidemiology of human influenza.

Once AI is established in domestic poultry, it is a highly contagious disease and wild birds are no longer an essential ingredient for spread. Infected birds excrete virus in high concentration in their faeces and also in nasal and ocular discharges. When introduced into a flock, the virus is spread from flock to flock by the usual methods involving the movement of infected birds, contaminated equipment, egg flats, feed trucks, and service crews, to mention a few. The disease generally spreads rapidly in a flock by direct contact, but on occasions spread is erratic.

Airborne transmission may occur if birds are in close proximity and with appropriate air movement. Birds are readily infected via instillation of virus into the conjunctival sac, nares, or the trachea. Some evidences indicate that virus can be recovered from the yolk and albumen of eggs laid by hens at the height of the disease. The possibility of vertical transmission is unresolved. In addition, the hatching of eggs from a diseased flock would likely be associated with considerable risk [5].

5 Clinical Features

Compared with the LPAI, the clinical signs of HPAI are severe and result in high mortality rates (can be 100%) in many species of birds, especially domestic fowl. Typically, after 3-7 days incubation time, the following clinical signs will emerge: Sudden death; Severe depression with ruffled feathers; Inappetence; Drastic decline in egg production; Edema of head and neck; Swollen and cyanotic combs and wattles; Excessive thirst; Watery diarrhea that begins as bright green and progresses to white; Swollen and congested conjunctiva with occasional hemorrhage; Diffuse hemorrhage between hocks and feet; Respiratory signs are dependent on tracheal involvement; Nasal and ocular discharge; Mucus accumulation (varies); Lack of energy; Coughing/sneezing; Incoordination; Nervous system signs such as paralysis. Death may occur prior to any symptoms or as late as a week after symptoms, though it is frequently within 48 hours [6].

6 Diagnosis

According to the clinical manifestations and the characteristics of epidemiology, the preliminary

diagnosis can be easily performed. However, the final conclusion cannot be drawn without the further Laboratory Diagnosis, for example, the isolation and identification of the causative virus. Commercially available type A influenza antigen-capture enzyme linked immunosorbent assay kits designed for use in human influenza have recently shown promise as a possible rapid diagnostic test for poultry. Furthermore, the molecular biological technique, i.e., PCR and molecular pathology will be useful for common and Differential Diagnosis. However, all the diagnosis procedures should be performed by specialists and in special institutions due to the potential risk.

7 Prevention and Treatment

From the views of epidemic angle, to control the source of infection; to cut the transmission pathway; and to protect the susceptible groups are still very important and efficient for AI prevention.

Inactivated oil-emulsion vaccines and viable vaccines prepared using naturally avirulent or attenuated strains can be vaccine candidates, the former is safer that the latter, even though the former is more expensive that the latter in cost. Some drugs, such as amantadine hydrochloride can be administered in drinking water to reduce disease losses, but drug-resistant viruses quickly emerged, negating the initial beneficial effects to some extent [3]. Other prevention policies and treatments remain under investigated, and we are looking forward to hearing good news as soon as possible.

Correspondence to:
Xiaofeng Ren
Department of Preventive Veterinary, College of Veterinary Medicine, Northeast Agriculture University
59 Mucai Street, Harbin, Heilongjiang 150030, China
Telephone: 0086-451-55190385
Email: rxfemail@yahoo.com.cn

References

[1] Katz JM. The impact of avian influenza viruses on public health. Avian Dis 2003;47 (3 Suppl):914-920.
[2] http://www.cdc.gov/flu/avian/outbreaks/asia.htm
[3] http://www.vet.uga.edu/vpp/gray_book/FAD/avi.htm.
[4] Alexander DJ. Should we change the definition of avian influenza for eradication purposes? Avian Dis 2003;47(3 Suppl):976-981.
[5] http://www.fao.org/ag/againfo/subjects/en/health/diseases-cards/avian.html
[6] http://www.cidrap.umn.edu/cidrap/content/biosecurity/ag-biosec/anim-disease/avianflu.html.

Nature and Science

Call for Papers

The new international academic journal, "**Nature and Science"** (ISSN: 1545-0740)**,** is registered in the United States, and invites you to publish your papers.

Any valuable papers that describe natural phenomena and existence or any reports that convey scientific research and pursuit are welcome, including both natural and social sciences. Papers submitted could be reviews, objective descriptions, research reports, opinions/debates, news, letters, and other types of writings that are nature and science related.

This journal will be supported by manuscript contributors. To cover the printing cost, this journal will only charge authors the actual printing fee (about US$30 per printed page). At least one hard copy of the printed journal will be given to each author free of charge. Here is a new avenue to publish your outstanding reports and ideas. Please also help spread this to your colleagues and friends and invite them to contribute papers to the journal. Let's work together to disseminate our research results and our opinions.

Papers in all fields are welcome, including articles of natural science and social science.

Please send your manuscript to editor@sciencepub.net.

For more information, please visit http://www.sciencepub.org.

Marsland Company
P.O. Box 21126
East Lansing, Michigan 48909
The United States
Telephone: (517) 980-4106
E-mail: editor@sciencepub.net
Website: http://www.sciencepub.org

Marsland Company

P.O. Box 753

East Lansing,Michigan 48826

The United States

Tel:(517)862 - 6881

http://www.sciencepub.org

E-mail:editor@sciencepub.net