ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of Advanced Research in Artificial Intelligence

Volume 4 Issue 10

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

Editorial Preface

From the Desk of Managing Editor…

Artificial Intelligence is hardly a new idea. Human likenesses, with the ability to act as human, dates back to Geek mythology with Pygmalion's ivory statue or the bronze robot of Hephaestus. However, with innovations in the technological world, AI is undergoing a renaissance that is giving way to new channels of creativity.

The study and pursuit of creating artificial intelligence is more than designing a system that can beat grand masters at chess or win endless rounds of Jeopardy!. Instead, the journey of discovery has more real-life applications than could be expected. While it may seem like it is out of a science fiction novel, work in the field of AI can be used to perfect face recognition software or be used to design a fully functioning neural network.

At the International Journal of Advanced Research in Artificial Intelligence, we strive to disseminate proposals for new ways of looking at problems related to AI. This includes being able to provide demonstrations of effectiveness in this field. We also look for papers that have real-life applications complete with descriptions of scenarios, solutions, and in-depth evaluations of the techniques being utilized.

Our mission is to be one of the most respected publications in the field and engage in the ubiquitous spread of knowledge with effectiveness to a wide audience. It is why all of articles are open access and available view at any time.

IJARAI strives to include articles of both research and innovative applications of AI from all over the world. It is our goal to bring together researchers, professors, and students to share ideas, problems, and solution relating to artificial intelligence and application with its convergence strategies. We would like to express our gratitude to all authors, whose research results have been published in our journal, as well as our referees for their in-depth evaluations.

We hope that this journal will inspire and educate. For those who may be enticed to submit papers, thank you for sharing your wisdom.

Editor-in-Chief
IJARAI
Volume 4 Issue 10 October 2015
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

Diagrammatic Representation as a Tool for Clarifying Logical Arguments

Sabah Al-Fedaghi
Computer Engineering Department
Kuwait University
Kuwait

***Abstract*—Knowledge representation of reasoning processes is a central notion in the field of artificial intelligence, especially for knowledge-based agents, because such representation facilitates knowledge of action outcomes necessary for optimum performance by problem-solving agents in complex situations. Logic is the primary vehicle by which knowledge is represented in knowledge-based agents. It involves logical inference that produces answers from what is known based on this inference mechanism. Modus Ponens is the best-known rule of inference that is sound. Recently, a dispute has arisen regarding attempts to show that modus ponens is not a valid form of inference. Part of the cause of the controversy is miscommunication of the involved problem. This paper proposes a diagrammatic representation of modus ponens with the hope that such a representation will serve to clarify the issue. The advantage of this diagrammatic representation is a better understanding of the reasoning process behind this inference rule.**

Keywords—artificial intelligence; diagrammatic representation; conditionals; argument forms; logical argumentation; modus ponens

I. INTRODUCTION

This paper is concerned with the representation of knowledge and the reasoning process, which are central notions in artificial intelligence, especially for knowledge-based agents. The subject is important for artificial agents because they facilitate knowledge of action outcomes necessary for their optimum performance in complex situations and partially observable environments.

Logic is a primary vehicle for representing knowledge. It involves logical inference that produces answers from what is known based on this inference mechanism. In addition to this function of reasoning, an inference mechanism can be utilized for self-learning by artificial agents. Knowledge bases founded on logic consist of statements that accept truth-values with respect to each possible world. They also involve logical *entailment* between statements, where statements follow logically from other statements. Entailment can be used to derive conclusions—that is, to carry out logical inference. An inference rule that derives only entailed sentences is said to be sound or truth preserving.

Modus Ponens (MP) is the best-known rule of inference that is sound or truth preserving and hence can be applied to derive conclusions that lead to the desired goal. History-wise, it has been considered one of the five basic inference rules that are valid without proof (e.g., by the Stoics [1]). Currently, it is still a central tool; for example, MP is an important forward-chaining inference in a knowledge base of Horn clauses to determine whether a statement is entailed by the knowledge base.

This paper focuses on such MP because recent attempts have been made to show that modus ponens is not a valid form of inference. The paper does not counter or support the logical argument of such an attempt. The underlying thesis is that representation has contributed to such a controversy. This paper aims at proposing a diagrammatic representation of modus ponens with the hope of clarifying the issue in relation to MP.

Current methods of diagrammatic representation of logic formulas incompletely depict the underlying semantics of the formulas, creating a conceptual gap that sometimes causes misinterpretation. The methodology proposed in this paper applies a schematizing of logic formulas in the context of modus ponens. The advantage of this diagrammatic representation is better understanding of the reasoning process at the base of this inference rule.

II. BACKGROUND OF MODUS PONENS

In logic, an indicative conditional statement is a statement that describes implications or hypothetical situations and their consequence, such as *If p then q*, where *p* is called the antecedent, and *q* the consequent; however, in general, such as in natural language, conditional statements are not restricted to this format. In the context of logic, and based on truth conditions, *If p then q*, denoted as $p \rightarrow q$, is false when p is true and q is false, otherwise, it is true. “Conditional sentences have attracted concentrated attention of philosophers, although intermittent, since ancient times…” [2].

On the other hand, MP as a principle of inference expresses that: from the conditional if p then q together with its antecedent p, it can be inferred that q. MP is commonly recognized as a basic rule of inference. Along with MP is the Modus Tollens (MT) rule: “we teach them [MP and MT] in introductory logic courses, related to conditional statements. In everyday reasoning, MP and MT can also have important roles, in modes of argumentation” [2]. As mentioned, attempts have been made to show that MP is not a valid form of inference, and these have been based mainly on a number of counterexamples, thus challenging the accepted view in logic that inferences grounded on MP are deductively valid [3-4]. McGee [5] presents the following MP counterexamples:

Opinion polls taken just before the 1980 election showed the Republican Ronald Reagan decisively ahead of the Democrat Jimmy Carter, with the other Republican in the race, John Anderson, a distant third. Those apprised of the poll results believed, with good reason:

a) If a Republican will win the election, then if Reagan will not win, Anderson will win.

b) A Republican will win the election.

c) So, if Reagan will not win, Anderson will win.

But, those apprised of the poll results "did not have reason to believe" conclusion c [5]. This means that c is not obviously true. This application of MP to an ordinary statement leads to a conclusion that is contrary to common-sense expectation. Accordingly, in light of examples such as this, modus ponens is not strictly valid; see discussions in [6-8]. "McGee's [5] attempt to show that modus ponens is not a valid form of inference – and to show this by the help of a counterexample and not by envisaging an evil demon confusing us – is proof of the ingenuity of a philosopher's ability to doubt" [9]. MacFarlane [10] gives two additional examples, as follows:

(a) If that creature is a fish, then if it has lungs, it is a lungfish. (b) That creature is a fish. (c) So, if it has lungs, it is a lungfish.

(a) If Uncle Otto doesn't find gold, then if he strikes it rich, he will strike it rich by finding silver. (b) Uncle Otto won't find gold. (c) So, if Uncle Otto strikes it rich, he will strike it rich by finding silver.

This paper demonstrates diagrammatic construction of MP for the purpose of producing a conceptually complete description of the involved phenomena. The description can provide illustrations and models that might help in facilitating understanding of the MP-based reasoning process. The approach utilizes a diagrammatic apparatus called the Flowthing Model that, for the sake of completeness, will be briefly described in the next section [11-15].

III. USING THE FLOWTHING MODEL

The Flowthing Model (FM) can be related to the notion of fluidity within a web of interrelated *flows* that cross boundaries of intersecting and nested *spheres*. This representation is an apparatus that facilitates flowages (acts of flowing). Ingredients in a flowage include *flowthings* (things that flow), and flow systems (*flowsystems*). So-called objects, concepts, entities, and time are flowthings. A "thing" is defined as a flowthing: "what is created, released, transferred, arrives, is accepted, and processed" while flowing within and among spheres. In spite of use of the term "thing," the fundamental ontology in FM is that "systems are not composed of things, but are rather *defined on* things, and there is a clear distinction between their physical 'thinghood' and logical 'systemhood' properties" [16]. Accordingly, a sphere or subsphere can be any object, any region of logical space that is set apart (mentally) from anything else [16].

A flowthing has a permanent identity but impermanent form. A *flowsystem* constrains the trajectory of flow of flowthings. A particular flowsystem provides the space/time for happenings and existence of flowthings. To flowthings, the flowsystem is formed of six discontinuities: being created, being released, being transferred, being arrived, being accepted, and being processed.

Flows connect six *stages* that are exclusive for flowthings; i.e., a flowthing can be in one and only one of these six states at a time: Transfer, Process, Creation, Release, Arrival, and Acceptance, as shown in Fig. 1. Where appropriate, we can use Receive as a combined stage of Arrive and Accept. These stages are the elementary basic *actions*. A system manifests itself by engaging in these actions: processing, creating, releasing, receiving, and transferring of flowthings. In Fig. 1, we assume irreversibility of flow, e.g., released flowthings flow only to Transfer.

Note that this conceptualization of stages as elementary *actions* may not coincide with other uses of such terms, e.g., in physics. For example, (model) time and (model) space are simply flowthings in FM that can be created, processed, released, etc.; e.g., a clock is a flowsystem that can create, release, and transfer time.

The lower-level *spheres* where the flows occur are called flowsystems; these include, at most, six stages, as follows:

- *Arrive*: a flowthing reaches a new flowsystem

- *Accepted*: a flowthing is permitted to enter the system.

- *Processed* (changed in form): the flowthing passes through some kind of transformation that changes its form but not its identity (e.g., compressed, colored, compared)

- *Released*: a flowthing is marked as ready to be transferred (e.g., airline passengers waiting to board after completing processing)

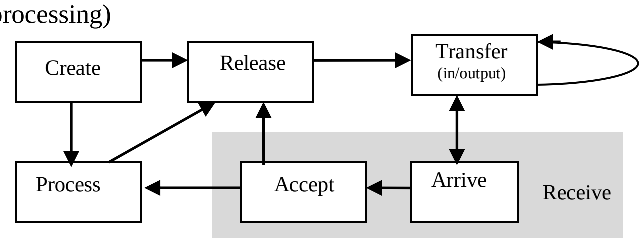


Fig. 1. Flow system

- *Created*: a new flowthing emerges (comes into existence relative to its sphere) in the system (e.g., processing of a neutron generates a proton, electron, and neutrino)

- *Transferred*: the flowthing is en route to somewhere outside the flowsystem (e.g., packets reaching ports in a router, but still not in the arrival buffer).

An additional stage of *Storage* can also be added to any FM model to represent the storage of flowthings; however, storage is a generic stage, not specific, because there can be stored processed flowthings, stored created flowthings, and so on.

A flowsystem may not need to include all the stages because the other stages are irrelevant, have no impact, or are prohibited, e.g., an archiving (storage) system might use only the stages arrive, accept, release, and transfer. Multiple systems captured by FM can interact with each other by triggering interrelated events in their spheres and stages.

IV. Describing Formulas in FM

In FM, a formula p can be conceptualized as a sphere formed from two subsystems (Body, Truth), as shown in Fig. 2. Consider the statement *A Republican will win the election* as declared in a logical argument:

1) …
2) …
3) *p*

This indicates that the inference rule, say, MP, processed the premises and reached the conclusion that triggered the creation of p (circles a and b in Fig. 3). As a result of this creation "(3) p" appears in the chain of deduction with its two flowsystems of truth-value and body (c and d).

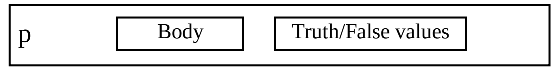


Fig. 2. p as a sphere with two flowsystems

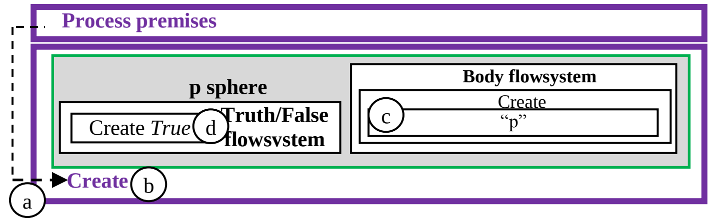


Fig. 3. p is created as a result of processing premises

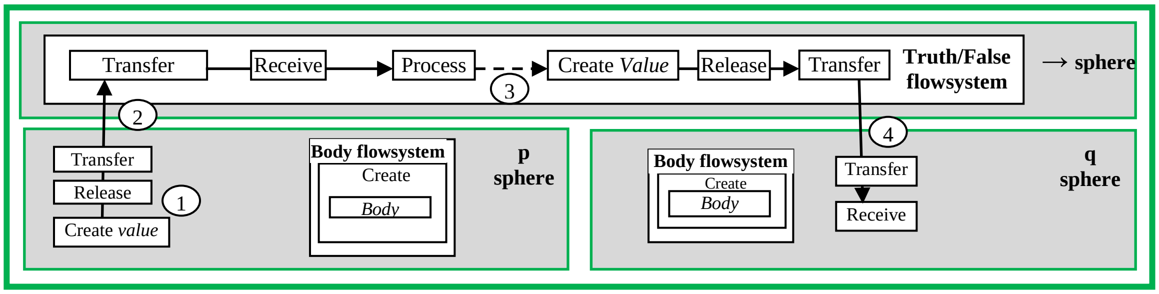


Fig. 4. FM representation of the implication ***p→q***

Fig. 4 shows the FM representation of the implication p→q which is formed from p', q', and →. For simplicity's sake, the truth-value flowsystem (circle 1 in the figure) is not enclosed in a box. The truth-value flows to p', then to the implication (2), where it is processed and, according to the material implication truth table, triggers (3) the creation of a truth-value. This truth-value flows to q' (4). Note that Fig. 4 is an "empty shell" of structure that will be filled when it is triggered. The implication includes p' and q' as shells (place holders) of structure (no assigned truth-values).

Now consider that the modus ponens:

4) *p→q*
5) *p*

is applied to produce q as shown in Fig. 5. Note that the sphere of the MP involves p, p→q, and q. Such a structure of MP (see Fig. 5) is activated (created) and processed. The antecedent and consequent p' and q' in the figure are "shells" or "place holders" for p and q (loaded with truth-values), in the same way one could give a value to variable *x* in *x* + 200. When p is created then, if its body is similar to the p', the truth-value flows and reaches → to trigger filling of q'; accordingly, the MP "gives birth to" (creates) q.

Fig. 6 shows the complete FM representation of MP, which involves the following:

a) The creation of p→q (shells p' and q', and →), p and q (1, 2, and 3, respectively). Note that true is assigned to the → sphere (2), activating it, analogous to switching an engine ON, as shown in Fig. 7.

b) Assigning a truth-value to p (1)

Accordingly, the *body* of p flows to its corresponding body of the antecedent in the implication (5). If the two bodies are identical (6), then this triggers (7) the flow of truth-value (8 and 9) to the implication to be processed (10) according to the implication truth table.

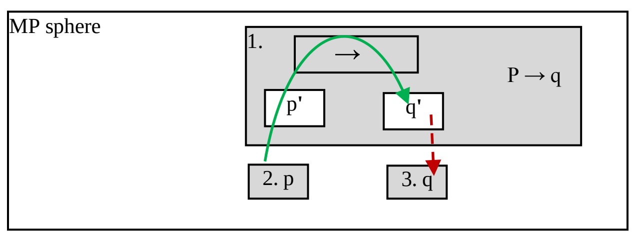


Fig. 5. The sphere of the MP

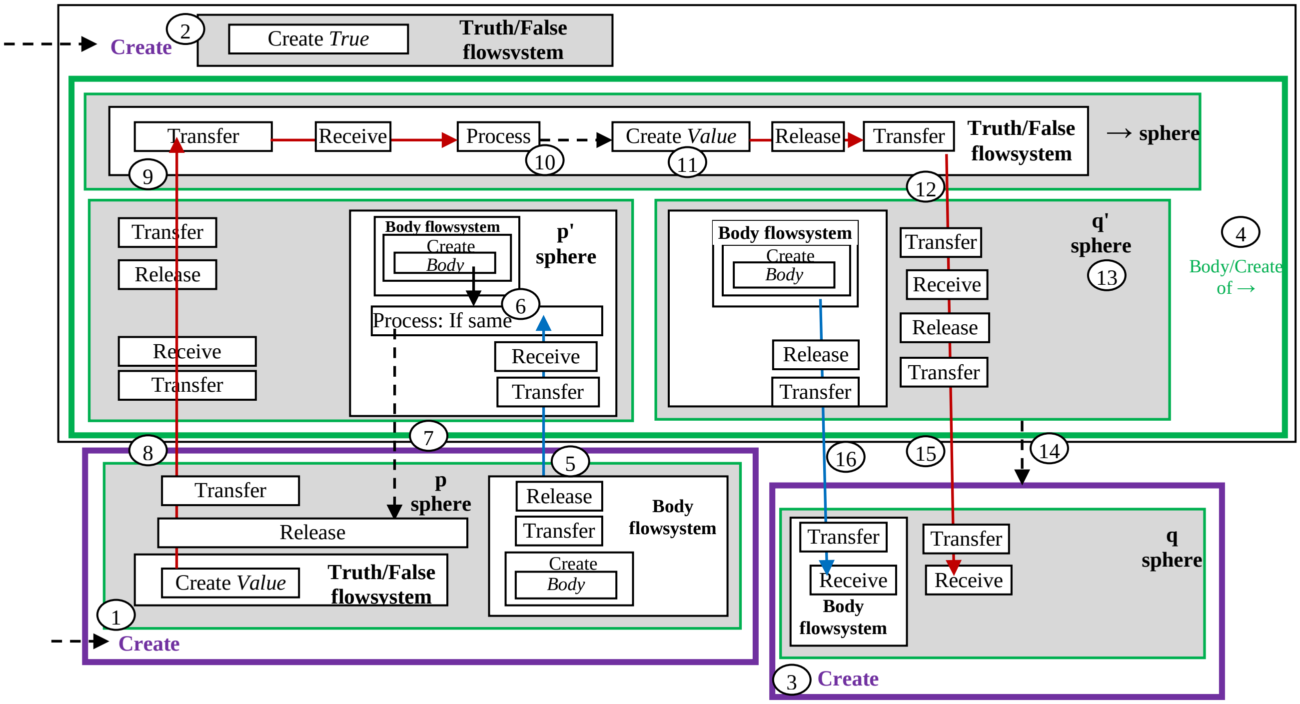


Fig. 6. FM representation of the MP: p→q, p that produces the conclusion q

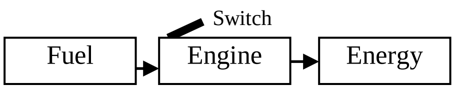


Fig. 7. Analogy of the activation of MP

The resultant truth-value is created (11) to flow (12) to the sphere of q' (13). This flow to the consequent causes it to trigger (14) the creation of q while "filling" it with a truth-value (15) and body (16).

V. Lewis Carroll

This section shows an example using the FM representation to clarify the mechanism of the modus ponens by drawing it explicitly.

"What the Tortoise Said to Achilles" was written by Lewis Carroll in 1895 as a regress problem that arises from using MP as a deduction rule. It begins by considering the following logical argument:

A: "Things that are equal to the same are equal to each other"

B: "The two sides of this triangle are things that are equal to the same"

Therefore Z: "The two sides of this triangle are equal to each other"

Then, an objection is raised to deducing Z from A and B, based on accepting that A and B are true, but not accepting the principle: *if* A and B are both true, *then* Z must be true. Accordingly, the premises are written as follows.

A: "Things that are equal to the same are equal to each other"

B: "The two sides of this triangle are things that are equal to the same"

C: "If A and B are true, Z must be true"

Therefore Z: "The two sides of this triangle are equal to each other"

However, it is possible to accept premise C while still refusing to accept the expanded argument. In this way, the list of premises continues to grow without end.

(1): "Things that are equal to the same are equal to each other"
(2): "The two sides of this triangle are things that are equal to the same"
(3): (1) and (2) ⇒ (Z)
(4): (1) and (2) and (3) ⇒ (Z)
...
(n): (1) and (2) and (3) and (4) and ... and (n − 1) ⇒ (Z)
Therefore (Z): "The two sides of this triangle are equal to each other."

Fig. 8 shows the FM representation of A, B, and Z. Now consider in the figure:

<u>Refusing to deduce Z from A and B based on accepting that A and B are true, but not accepting the principle: *if* A and B are both true, *then* Z must be true.</u>

The principle that, *if* A and B are both true, *then* Z must be true, is drawn explicitly in the figure as an application of Fig. 6. It seems that refusal is related to the triggering that creates q (*The two sides of this triangle are equal to each other*).

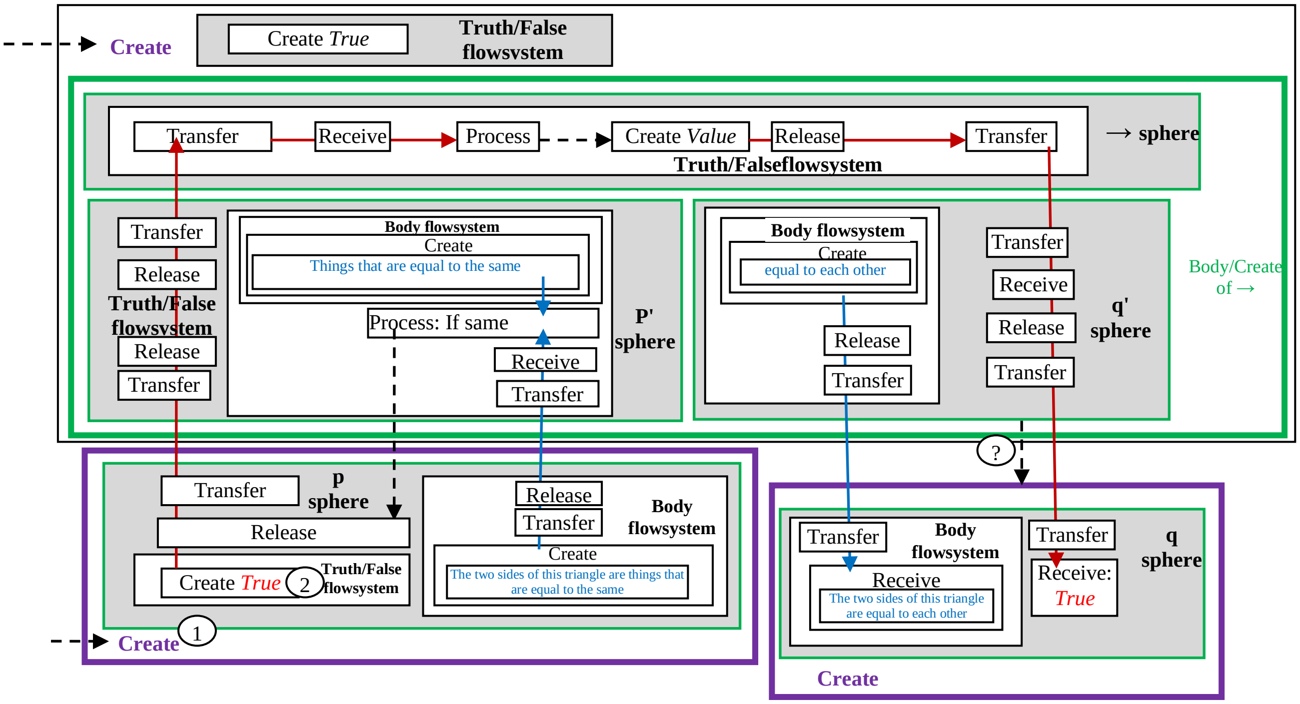


Fig. 8. FM representation of the MP: A: "Things that are equal to the same are equal to each other," B: "The two sides of this triangle are things that are equal to the same"; therefore Z: "The two sides of this triangle are equal to each other"

As explained when discussing p→q in Fig. 6, p is created (given), and its truth-value (2) flows to the implication → where the truth-value is created according to the truth table (3). Hence, q' is now "pregnant" with full q: *It is true that the two sides of this triangle are equal to each other*. Then, why disbelieve that q' "gives birth" to q? The whole process is a machine-like construction analogous to a machine designed to produce an output.

Another possible objection is disbelieving that the "machine" is designed correctly. What part, then, is the incorrect portion of the machine?

FM representation allows the mechanism of the modus ponens to be explicitly drawn, in contrast to being a "ghost" in such representations as the one shown in Fig. 9.

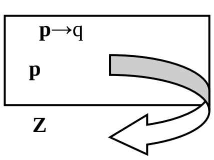


Fig. 9. Implicit representation of Modus ponens

VI. McGee's Counterexample

The FM representation can be used to diagram McGee's [5] counterexample mentioned in the introduction.

a) If a Republicans will win the election, then if Reagan will not win, Anderson will win.

b) A Republican will win the election.

c) So, if Reagan will not win, Anderson will win.

Or,

1. p→(q →t)
2. p
3. (q →t)

Fig. 10 shows the corresponding FM representation. The first part, "p→" appears as in the FM description shown in Fig. 6; however, starting with circle 12, the truth-value result, this time, triggers (activates) *another implication* (blue box in the online version): (q →t).

This, in turn, triggers the creation of q (circle 13 in Fig. 10), which is formulated from the body of q' in the implication (14 and 15).

Accordingly, the *true* value of q' (16) flows to the implication in (q' →t') (17), where it is processed (18) to produce a truth-value according to the material implication truth table (19). Since (q' →t') is true and q is true, then the generated truth-value is true. This truth-value flows to (the shell) t' (20) to trigger (21) the creation of t using the body of t' (22).

Now, look at the "internal" MP:

q→t
q
----
t

The situation of (q'→t') (12) being true does not necessarily originate from q is true, as shown in Table1.

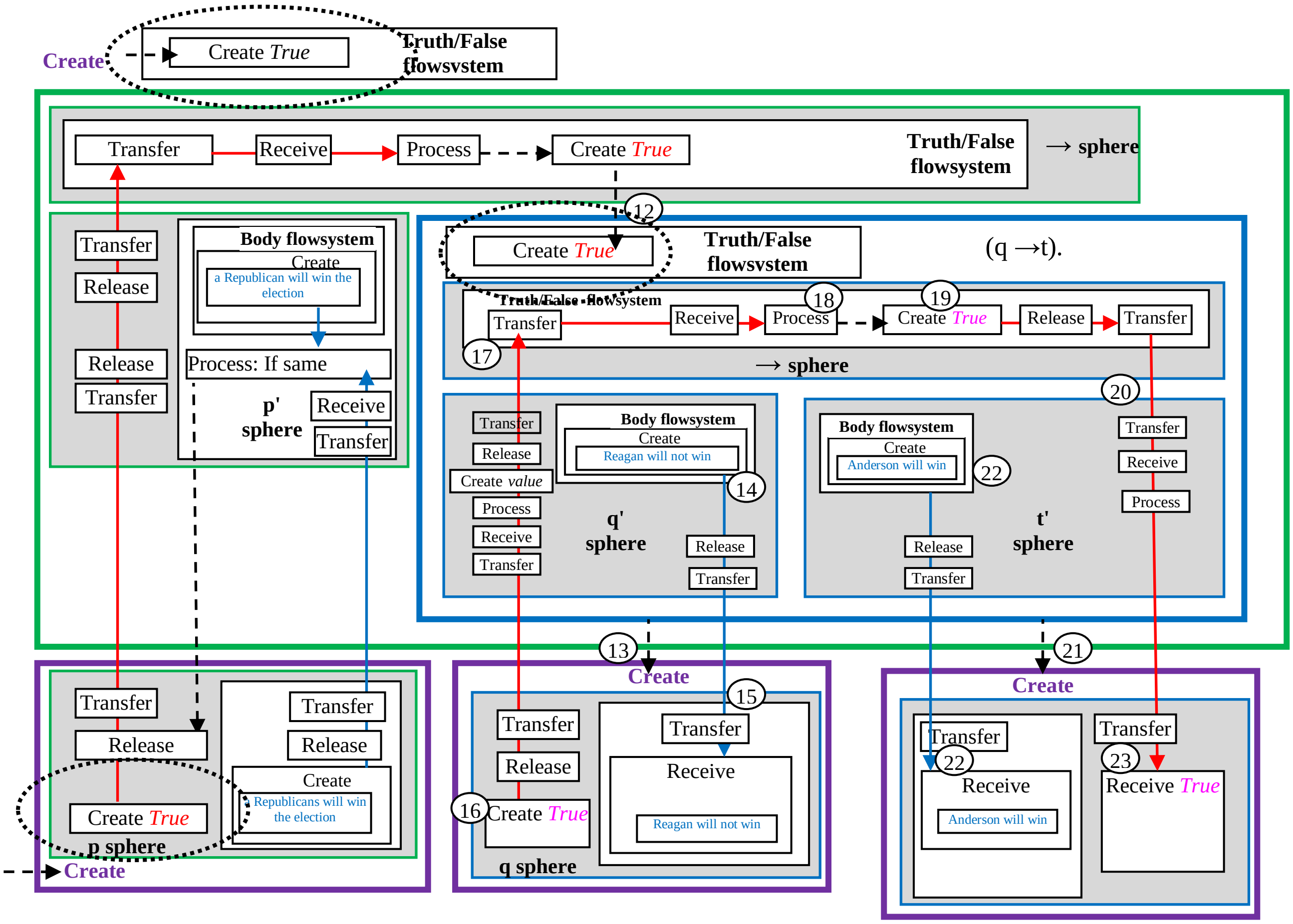


Fig. 10. FM representation of McGee's counterexample

TABLE I. TRUTH TABLE FOR "q→t" IS TRUE"

	q	t	(q →t)
1	**true**	**true**	**true**
2	false	false	true
3	false	true	true

The true value of (q→t) "means" that q and t could be true or false, as shown in Table 1, relative to the assumed truthfulness of *Reagan will not win*. Accordingly, the implication q→t has three possibilities, all of which satisfy that (q →t) is true:

(Row 1 in Table 1) *A Republican will win the election → If Reagan will not win, Anderson will win*

(Row 1 in Table 1) *A Republican will win the election → If Reagan will win, Anderson will win*

(Row 1 in Table 1) *A Republican will win the election → If Reagan will win, Anderson will not win*

Therefore, the MP should have been written as:

1) If a Republican will win the election, then (If Reagan will not win, Anderson will win) ∨ (If Reagan will win, Anderson will win) ∨ (If Reagan will win, Anderson will not win)

2) A Republican will win the election

----

3) (If Reagan will not win, Anderson will win) ∨ (If Reagan will win, Anderson will win) ∨ (If Reagan will win, Anderson will not win)

But for all $p_1 \vee p_2 \vee p_3$, $p_1 \vee p_2 \vee p_3$ is true if any of p_1, p_2, or p_3 is true. Accordingly, the consequent (3) is true because (*If Reagan will win, Anderson will not win*). In general, if p_i is true then ($p_i \vee$ any false statement) is true. The controversy originated with the implication:

p is true → ((q→t) is true)

Subsequently, we can substitute a false statement for q and t and still preserve the truthfulness. If fact, it is a valid deduction that:

1) If a Republican will win the election, then, if The moon is made of green cheese, Anderson will win

2) A Republican will win the election

----

3) (If The moon is made of green cheese, Anderson will win)

This resulted from the definition of material implication [16].

VII. Conclusion

This paper proposes a diagrammatic representation of modus ponens with the hope that such a representation can help to clarify issues related to rules of inference, specifically modus ponens. The advantage of this diagram-matic representation as a tool for understanding the reasoning process involved in this inference rule is demonstrated through examples. The results point to the viability of the approach. Further research may confirm such results.

References

[1] B. Mates, Stoic Logic. Berkeley: University of California Press, 1953. ISBN 0-520-02368-4

[2] M. M. Dagli, "Modus ponens, modus tollens, and likeness," in Twentieth World Congress of Philosophy, August 10-15, Boston, USA, 1998.

[3] R. Reiter, "On reasoning by default," in Readings in Knowledge Representation, R. J. Brachman and H. J. Levesque, Eds. Los Altos, CA: Morgan Kaufman, 1985, pp. 402-410.

[4] H. Prakken, Logical Tools for Modelling Legal Arguments. Dordrecht: Kluwer, 1997.

[5] V. McGee, "A counterexample to modus ponens." J. Philos., vol. 82, pp. 462-471, 1985.

[6] W. Sinnott-Armstrong, J. Moor, and R. Fogelin, "A defense of modus ponens," J. Philos., vol. 83, pp. 296-300, 1986.

[7] E. J. Lowe, "Not a counterexample to modus ponens," Analysis, vol. 47, pp. 44-47, 1987.

[8] D. E. Over, "Assumptions and the supposed counterexamples to modus ponens," Analysis, vol. 47, pp. 142-146, 1987.

[9] C. Piller, "Vann McGee's counterexample to modus ponens," Philos. Stud., vol. 82, pp. 27-54, 1996.

[10] J. MacFarlane, "McGee on modus ponens," April 21, 2011. http://johnmacfarlane.net/142/mcgee.pdf

[11] S. Al-Fedaghi, "Schematizing proofs based on flow of truth values in logic," in IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2013), October 13-16, Manchester, UK, 2013.

[12] S. Al-Fedaghi, "Visualizing logical representation and reasoning," in 15th International Conference on Artificial Intelligence (ICAI'13), July 22-25, Las Vegas, USA, 2013.

[13] S. Al-Fedaghi, "Schematizing formulas for logic students," Int. J. Intell. Inform. Process., vol. 4, no. 4, pp. 27- 38, 2013.

[14] S. Al-Fedaghi, "On a flow-based paradigm in modeling and programming," Int. J. Advanced Comput. Sci. Appl. (IJACSA), vol. 6, no. 6, pp. 209-217, 2015.

[15] S. Al-Fedaghi, "Conceptualizing effects, and uses of information," in Information Seeking in Context Conference (ISIC 2008), September 17-20, Vilnius. Lithuania, 2008.

[16] G. Sundholm, "'Inference versus consequence' revisited: inference, consequence, conditional, implication," Synthese, vol. 187, pp. 943–956, 2012.

Relation Between Chlorophyll-A Concentration and Red Tide in the Intensive Study Area of the Ariake Sea, Japan in Winter Seasons by using MODIS Data

Kohei Arai 1
1Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Relation between chlorophyll-a concentration and red tide in the intensive study area of the back of Ariake Sea, Japan in the recent winter seasons is investigated by using MODIS data. Mechanism of red tide appearance is not so clarified. On the other hand, chlorophyll-a concentration can be estimated with satellite remote sensing data. An attempt is made for estimation of the location and size of red tide appearance. In particular, severe damage due to red tide is suspected in the winter seasons now a day. Therefore, 6 years (winter 2010 to winter 2015) data of MODIS data derived chlorophyll-a concentration and truth data of red tide appearance (the location and the volume) which are provided by Saga Prefectural Fishery Promotion Center: SPFPC (once/10 days of shipment data) have been investigated. As the results of the investigation, it is found that a strong correlation between the chlorophyll-a concentration and red tide appearance together with the possible sources of the red tide.**

Keywords—chlorophyl-a concentration; red tide; diatom; MODIS; satellite remote sensing

I. INTRODUCTION

The Ariake Sea is the largest productive area of Nori (Porphyra yezoensis[1]) in Japan. In winters of 2012, 2013, 2014 and 2015, a massive diatom bloom occurred in the Ariake Sea, Japan [1]. In case of above red tides, bloom causative was Eucampia zodiacus[2]. This bloom has being occurred several coastal areas in Japan and is well reported by Nishikawa et al. for Harimanada sea areas [2]-[10]. Diatom blooms have recurrently occurred from late autumn to early spring in the coastal waters of western Japan, such as the Ariake Sea [11] and the Seto Inland Sea [12], where large scale “Nori” aquaculture occurs. Diatom blooms have caused the exhaustion of nutrients in the water column during the “Nori” harvest season. The resultant lack of nutrients has suppressed the growth of “Nori” and lowered the quality of “Nori” products due to bleaching with the damage of the order of billions of yen [3].

This bloom had been firstly developed at the eastern part of the Ariake Sea. However, as the field observation is time-consuming, information on the developing process of the red tide, and horizontal distribution of the red tide has not yet been clarified in detail. To clarify the horizontal distribution of red tide, and its temporal change, remote sensing satellite data is quite useful.

In particular in winter, almost every year, relatively large size of diatoms of *Eucampia zodiacus* appears in Ariake Sea areas. That is one of the causes for damage of *Porphyra yezoensis*. There is, therefore, a strong demand to prevent the damage from Nori farmers. Since 2007, *Asteroplanus karianus* appears in the Ariake Sea almost every year. In addition, *Eucampia zodiacus* appears in Ariake Sea since 2012.

The chlorophyll-a concentration algorithm developed for MODIS[3] has been validated [13]. The algorithm is applied to MODIS data for a trend analysis of chlorophyll-a distribution in the Ariake Sea area in winter during from 2010 to 2015 is made [14]. On the other hand, red tide appearance (location, red tide species, the number of cells in unit water volume by using microscopy) are measured from the research vessel of the Saga Prefectural Fishery Promotion Center: SPFPC by once/10 days basis. The location and size of the red tide appearance together with the red tide source would be clarified by using SPFPC data. Match-up data of MODIS derived chlorophyll-a concentrations are used for investigation of relations between MODIS data and truth data of the red tide appearance.

In the next section, the method and procedure of the experimental study is described followed by experimental data and estimated results. Then conclusion is described with some discussions.

II. METHOD AND PROCEDURE

A. The Procedure

The purposes of the research is as follows,

1) At first, MODIS derived chlorophyll-a concentration has to be validated with the truth data (shipment data of the number of cells/ml of red tide species provided by SPFPC),

2) Possible sources of the red tide species has to be estimated,

[1] http://en.wikipedia.org/wiki/Porphyra

[2] http://www.eos.ubc.ca/research/phytoplankton/diatoms/centric/eucampia/e_zodiacus.html

[3] http://modis.gsfc.nasa.gov/

3) Spatial relation among several districted sea areas has to be clarified (more precisely, red tide relations between Ariake bay and Isahaya bay as well as Kumamoto offshore have to be clarified),

4) Mechanism of red tide in the intensive study sea areas will be clarified after all.

Therefore, the following procedure of the experimental study is proposed,

1) Gather MODIS data of the Ariake Sea areas together with the chlorophyll-a concentration estimation with the MODIS data,

2) Gather the truth data of red tide appearance (the location and the size of the red tide) together with the red tide species and the number of cells in unit water volume,

3) Investigation on relation between the truth data and the match-up of MODIS data.

B. The Intensive Study Areas

Fig.1 shows the intensive study areas in the Ariake Sea area, Kyushu, Japan.

Fig. 1. Intensive study areas

III. Experiments

A. The Data Used (MODIS Data Derived Chlorophyll-a Concentration and Truth Data)

MODIS derived chlorophyll-a concentration which area acquired for the observation period of one month (in January) in 2010 to 2015 is used for the experiments. MODIS data are acquired on these days. MODIS data cannot be acquired on the rest of days due to cloudy condition. White portions in the chlorophyll-a concentration images are cloud covered areas.

These data are acquired on January 4, 6, 7, 8, 9, 9[4], 10, 12, 17, 18, 20, 23, February 1, 3, 6, 9, 13, 14, 20, 27, and March 2 in 2015, respectively. Meanwhile, MODIS data are acquired on January 10, 13, 15, 16, 19, 23, 24, 26, 27, 29, 30 and February 4 in 2014, respectively. In 2013, MODIS data are acquired on January 4, 6, 10, 12, 15, 18, 19, 25, 28, 30, and 31, respectively while, in 2012. MODIS data are acquired on January 2, 6, 7, 12, 17, 20, 21, 23, 26, 29, 30, and 31, respectively. Furthermore, in 2011, MODIS data are acquired on January 1, 2, 7, 8, 14, 17, 22, 26, and 27, respectively while, in 2010, MODIS data are acquired on January 1, 3, 4, 9, 14, 16, 17, 18, 22, 24, 26, 27, and 29, respectively. All the data are shown in the previous paper which deals with "Locality of Chlorophyll-a Distribution in the Intensive Study Area of the Ariake Sea, Japan in Winter Seasons Based on Remote Sensing Satellite Data".

It is found the following red tide at around the Shiota river mouth on January 21 2010,

Asterionella kariana; 3280 cells/ml

Skeletonema costatum: 1330 cells/ml

Fig.2 shows the superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center. The number in the figure denotes the number of red tide cells / ml.

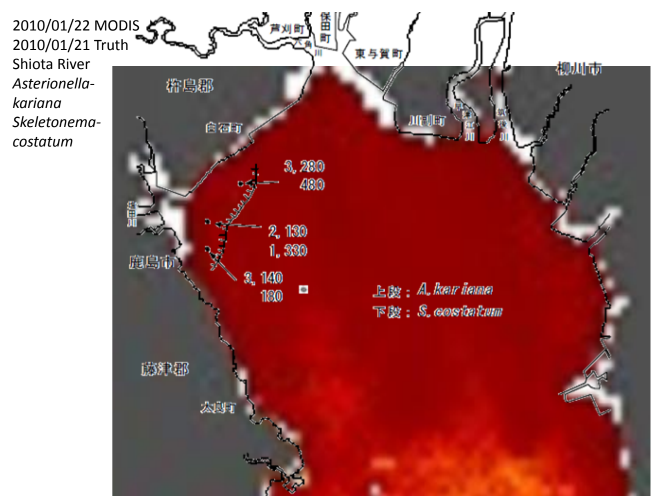


Fig. 2. Superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center

[4] There are two satellites which carry MODIS instruments, Terra and AQUA. Two MODIS data derived chlorophyll-a concentrations can be acquired occasionally.

As shown in the figure, the possible source of red tide would be nutrition rich water flown from the Shiota river mouth. MODIS derived chlorophyll-a concentration is distributed in the whole Ariake bay area while truth data shows red tide is distributed around the Shiota river mouth and Shiroishi offshore.

On January11 2011, it is found the following red tide along with the Shiroishi town offshore to the Shiota river mouth,

Asterionella kariana; 10150 cells/ml

Fig.3 shows the superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center. The number in the figure denotes the number of red tide cells / ml.

Although the truth data say that the red tide is distributed at around Shiota river mouth and Shiroishi offshore, it cannot be seen due to the fact that it is covered with cloud in the MODIS data derived chlorophyll-a concentration.

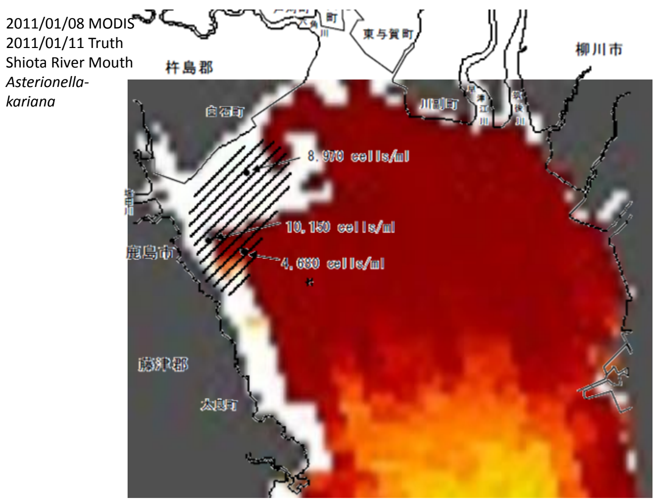


Fig. 3. Superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center

It is found the following red tide at around the Kashima offshore on February 25 2011,

Asterionella kariana; 4950 cells/ml

Fig.4 shows the truth data of red tide distribution which appeared at around Kashima offshore. Unfortunately, MODIS data cannot be acquired on that day.

It is found the following red tide at around the Shiota River Mouth on December 30 2011,

Asterionella kariana; 5150 cells/ml

Fig.5 shows the superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center.

The red tide is distributed Shiota river mouth and Kashima offshore.

On January 23 2012, it is found the following red tide at the Shiroishi offshore,

Skeletonema spp .: 5150 cells/ml

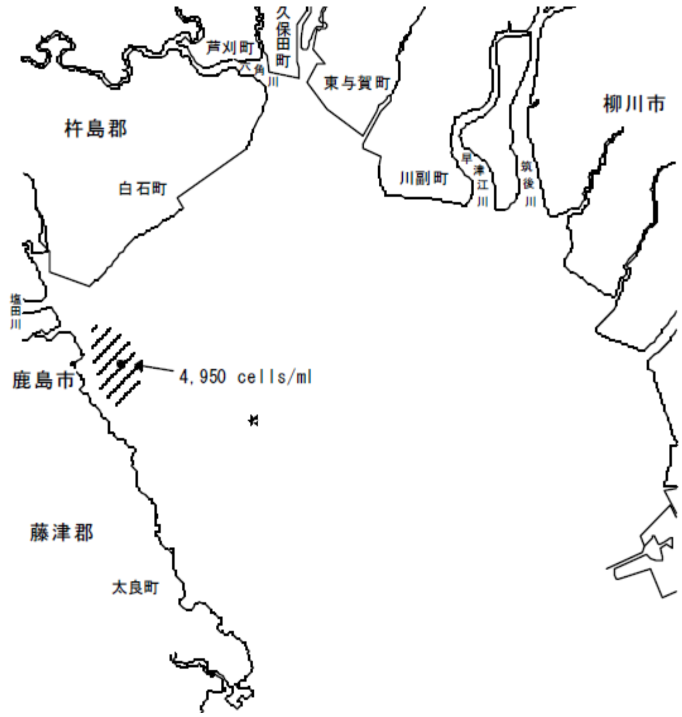


Fig. 4. Truth data of red tide distribution which appeared at around Kashima offshore

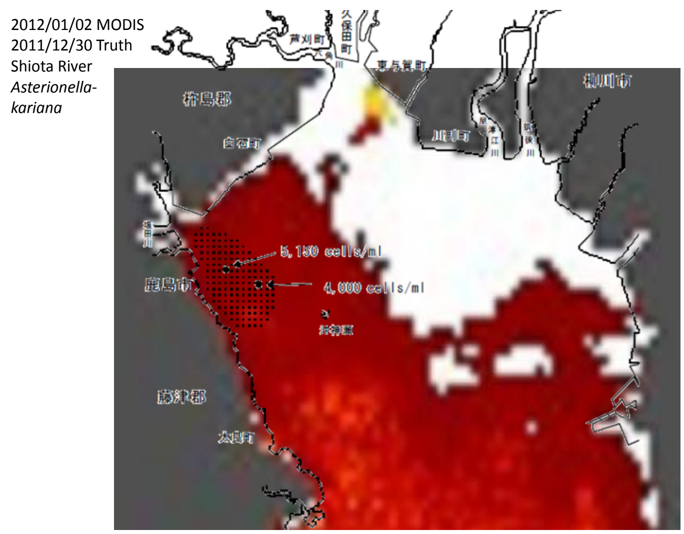


Fig. 5. Superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center

Fig.6 shows the superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center. The red tide is distributed at around Shiota river mouth and Shiroishi offshore.

The following red tide is found widely along with the Kawazoe offshore to the Tara offshore on February 22 2012,

Eucampia zodiacus: 1,090 cells/ml

Fig.7 shows the truth data of red tide distribution which appeared at around Kawazoe offshore. Unfortunately, MODIS data cannot be acquired on that day.

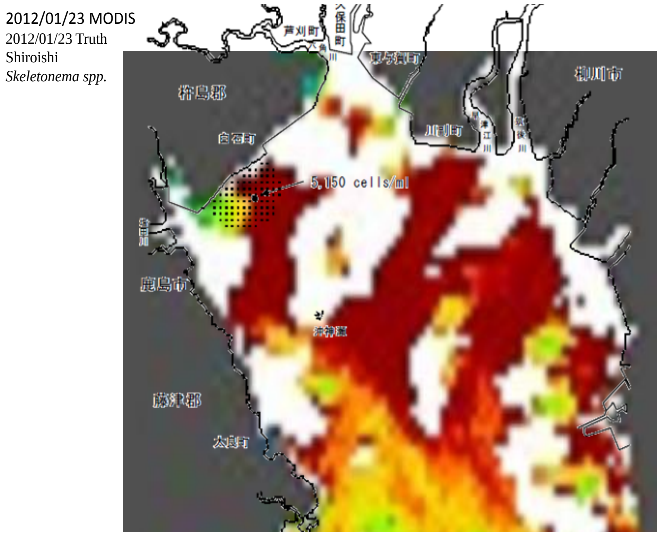


Fig. 6. Superimposed image with MODIS data derived chlorophyll-a concentration and truth data which is provided by Saga Prefectural Fishery Promotion Center

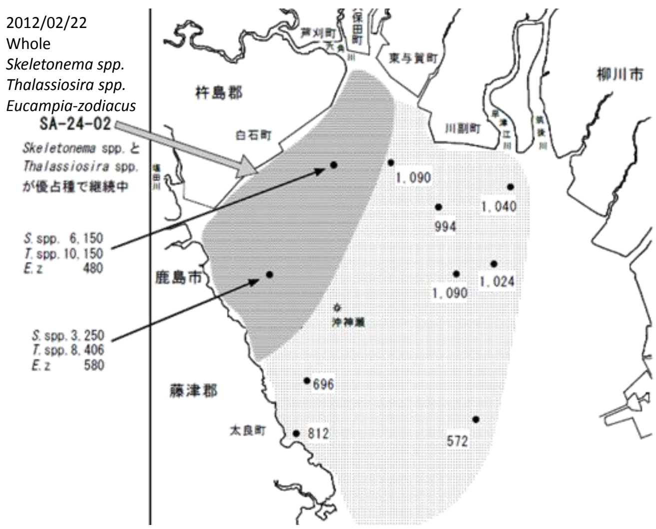


Fig. 7. Truth data of red tide distribution which appeared at around Kawazoe offshore

Also it is found the following red tide along with the Shiota river mouth to the Kashima offshore on December 31 2012,

Skeletonema spp .: 6110 cells/ml

Fig.8 shows the truth data of red tide distribution which appeared at around Kashima offshore and MODIS data derived chlorophyll-a concentration which is acquired on January 4 2013.

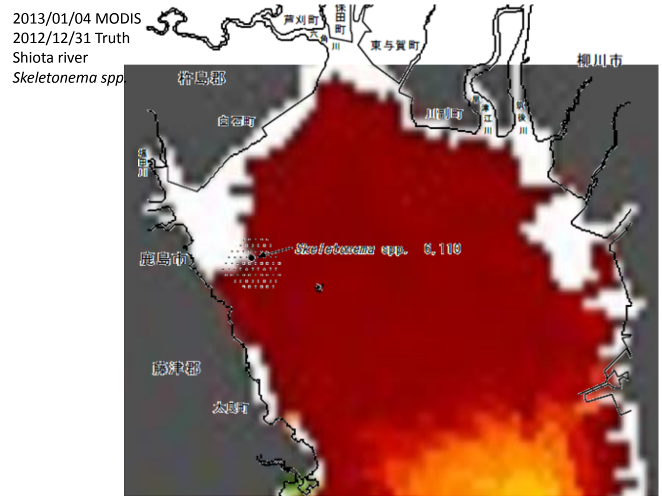


Fig. 8. Truth data of red tide distribution which appeared at around Kawazoe offshore and MODIS data derived chlorophyll-a concentration which is acquired on January 4 2013

On January 7 2013, the following red tide are observed along with the Shiota river mouth to the Shiroishi offshore,

Asterionella kariana; 5630 cells/ml

Skeletonema costatum: 3390 cells/ml

Fig.9 shows superimposed image of the truth data and the MODIS data derived chlorophyll-a concentration. The red tide distribution derived from MODIS data is almost coincident to the truth data.

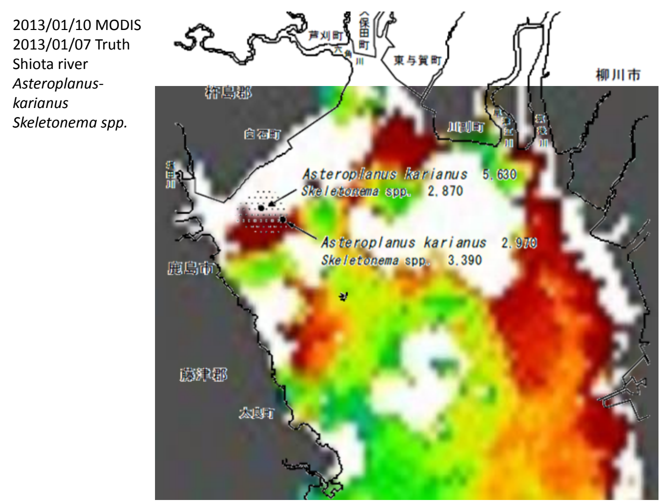


Fig. 9. Superimposed image of the truth data and the MODIS data derived chlorophyll-a concentration

It is found the following red tide at a small area of the Kashima offshore on February 18 2013,

Eutreptia pertyi and Eutreptiella spp.: 116600 cells/ml

Asterionella kariana; 7340 cells/ml

Fig.10 shows the truth data of red tide distribution which appeared at around Kashima offshore. Unfortunately, MODIS data cannot be acquired on that day.

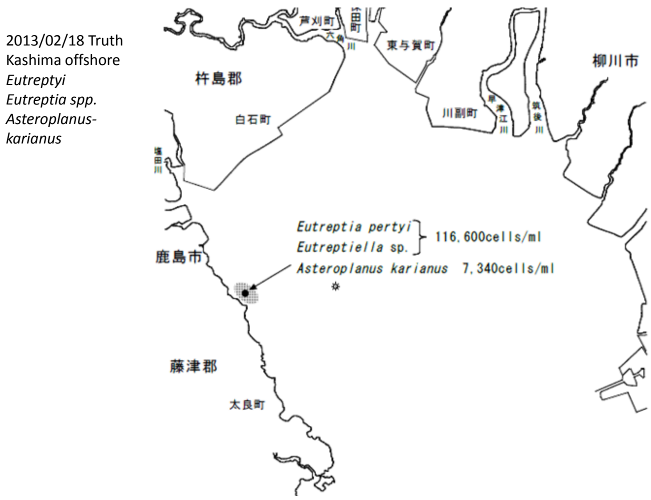


Fig. 10. Truth data of red tide distribution which appeared at around Kashima offshore

On February 26 2013, the following red tide is found almost entire Ariake Bay areas except Rokkaku river mouth,

Eucampia zodiacus: 980 cells/ml

Rhizosolenia setigera: 58 cells/ml

Fig.11 shows the truth data of red tide distribution. Unfortunately, MODIS data cannot be acquired on that day.

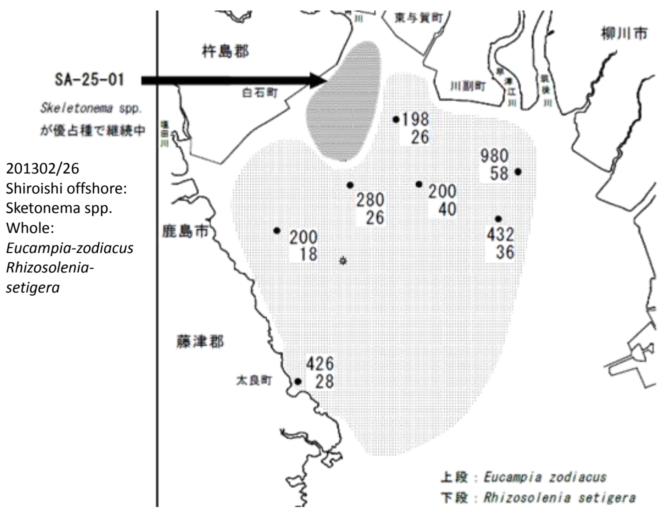


Fig. 11. Truth data of red tide distribution (almost all over the Ariake Bay area)

It is found the following red tide at the Shiraishi offshore on January 6 2014,

Asterionella kariana; 4830 cells/ml

Fig.12 shows the superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 10 2014.

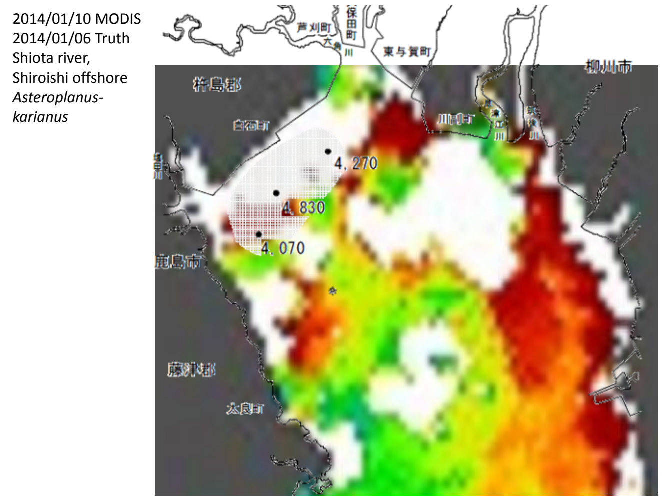


Fig. 12. Superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 10 2014

The following red tide is observed at the Shiroishi offshore on January 16 2014,

Skeletonema spp .: 6110 cells/ml

Thalassiosira spp.: 1510 cells/ml

Fig.13 shows the superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 16 2014. It seems that the red tide which is originated from Rokkaku river mouth and Shiota river mouth propagated to Shiroishi offshore along with the sea water current.

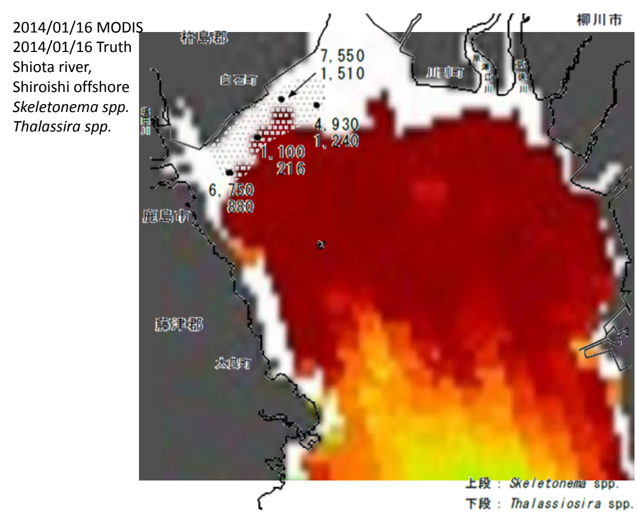


Fig. 13. Superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 16 2014

On February 6 2014, the following red tide is observed almost whole Ariake bay area except the Shiraishi offshore,

Eucampia zodiacus: 568 cells/ml

Fig.14 shows the superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on February 4 2014.

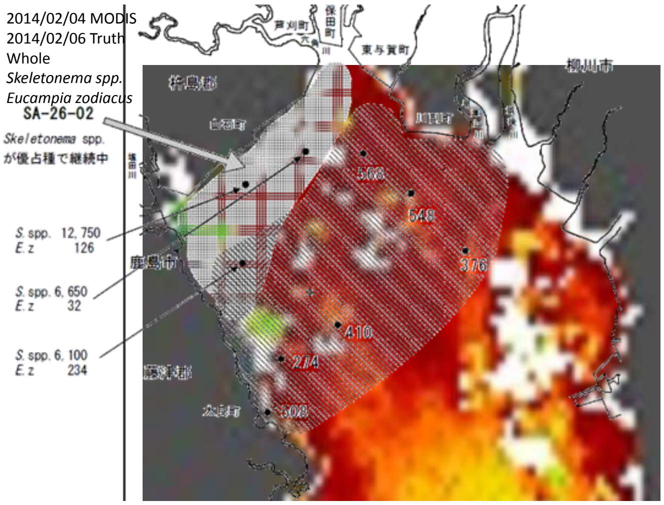


Fig. 14. Superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on February 4 2014

It is observed the following red tide along with the Shiroishi offshore to the Tara offshore on December 30 2014,

Asterionella kariana; 3890 cells/ml

Skeletonema costatum: 8750 cells/ml

On the other hand, MODIS data is acquired on January 4 2015, clouds are observed almost everywhere in the Ariake bay area though. Fig.15 shows the super imposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 4 2015. It seems that the red tide which is originated from the Shiota river mouth propagated to the Kashima and Tara offshore and a far beyond along with the sea water current.

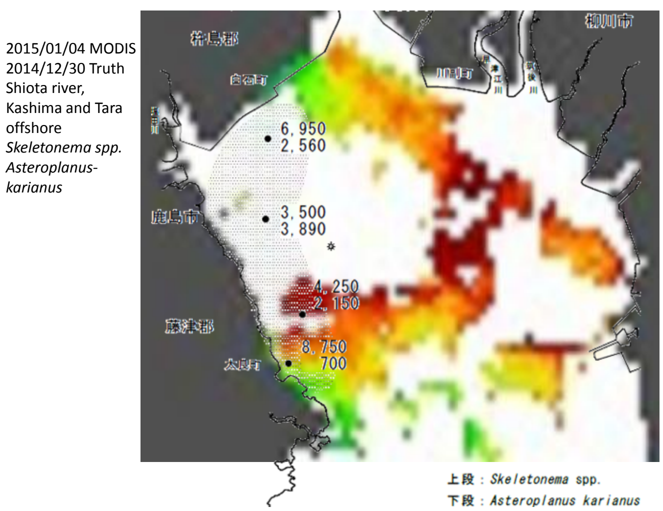


Fig. 15. Superimposed image of the truth data of red tide and the MODIS data derived chlorophyll-a concentration which is acquired on January 4 2015

On March 6 2015, the following red tide is observed along with the Kashima offshore to the Tara offshore,

Eucampia zodiacus: 1310 cells/ml

Fig.16 shows the superimposed image of truth data and the MODIS data derived chlorophyll-a concentration which is acquired on March 5 2015.

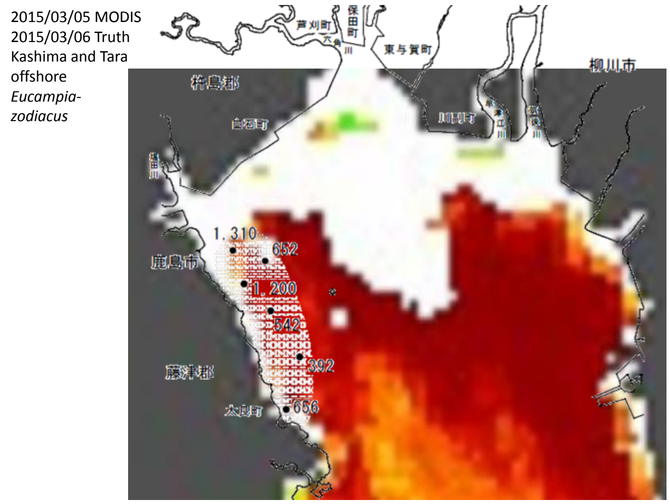
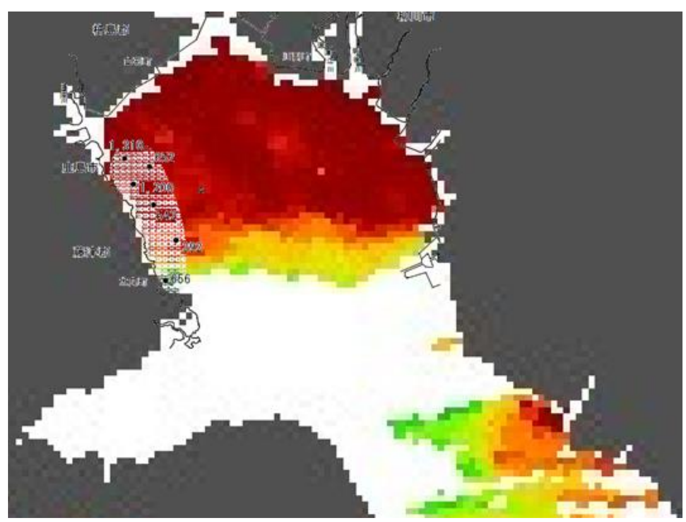


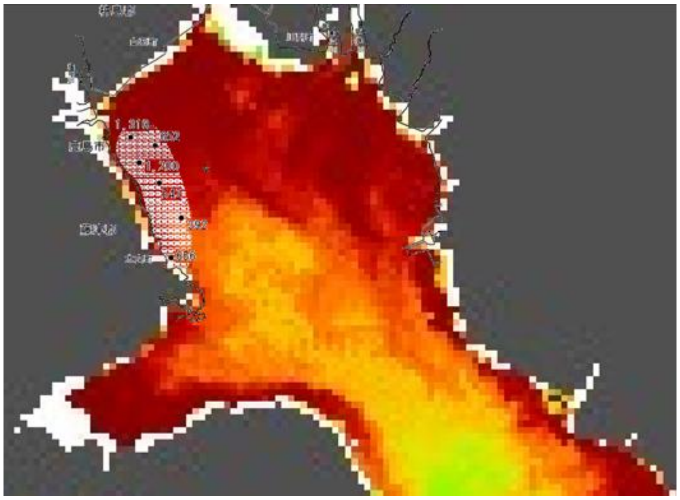
Fig. 16. Superimposed image of truth data and the MODIS data derived chlorophyll-a concentration which is acquired on March 5 2015

B. Trend Analysis in the Case of March 2015

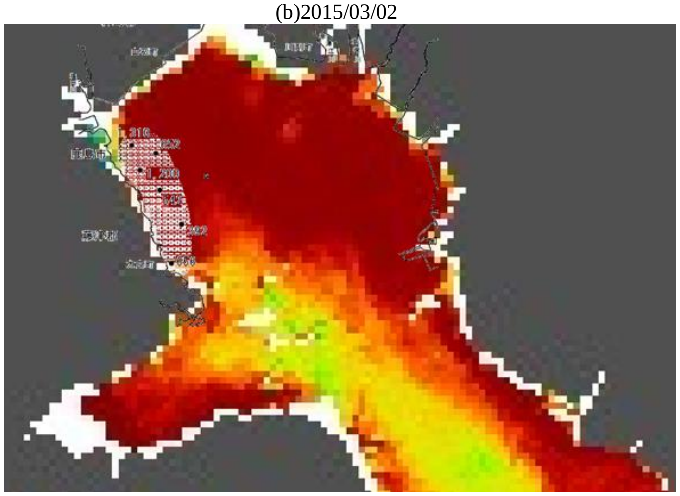
MODIS observed the same sea areas more frequently in this period. Fig.17 shows the superimposed images of the truth data and the MODIS data derived chlorophyll-a concentration which are acquired in the period starting from February 27 to March 5 2015. Chlorophyll-a is distributed densely in the Ariake bay area and Isahaya bay area on February 27. Then the densely distributed chlorophyll-a is flown to the south direction along with the sea water current in the Ariake bay while the densely distributed chlorophyll-a is flown from the Isahaya bay to the Taira-machi and far beyond the Shimabara offshore. Therefore, it may say that the sources of red tide are different between Ariake bay and Isahaya bay.



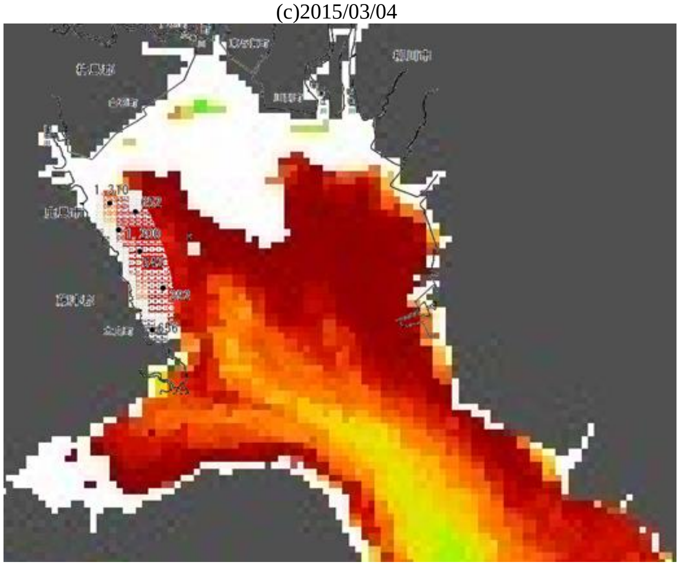
(a)2015/02/27



(b)2015/03/02



(c)2015/03/04



(d)2015/03/05

Fig. 17. Superimposed images of the truth data and the MODIS data derived chlorophyll-a concentration

IV. Conclusion

Through experiments with the MODIS data derived chlorophyll-a concentration and the truth data of red tide (Species and the number of cells/ml) which is provided by the Saga Prefectural Fisher Promotion Center, those are acquired for 6 years (winter 2010 to winter 2015), it is found the followings,

1) Asterionella kariana and Skeletonema costatum are used to be appeared in the Ariake Bay area in the winter seasons followed by Eucampia zodiacus appeared in the early spring every year after 2012 in particular, on February 22 2012, February 26 2013, February 6 2014 and March 6 2015.

2) It seems that the source of Asterionella kariana and Skeletonema costatum are mostly originated from Shiota river mouth and sometime from Rokkaku river mouth.

3) The red tide propagate from the center of Ariake Bay to Kashima, Tara offshore along with the sea water current in the counter clock wise direction.

4) Through the trend analysis with the superimposed images of the truth data and the MODIS data derived chlorophyll-a concentration which are acquired in the period starting from February 27 to March 5 2015, it is found that chlorophyll-a is distributed densely in the Ariake bay area and Isahaya bay area on February 27. Then the densely distributed chlorophyll-a is flown to the south direction along with the sea water current in the Ariake bay while the densely distributed chlorophyll-a is flown from the Isahaya bay to the Taira-machi and far beyond the Shimabara offshore. Therefore, it may say that the sources of red tide are different between Ariake bay and Isahaya bay.

Further investigations are required to clarify the mechanism of red tide appearance with the consideration three dimensional of cross section analysis the red tide source movement.

Acknowledgment

The authors would like to thank Dr. Toshiya Katano of Tokyo University of Marine Science and Technology, Dr. Yuichi Hayami, Dr. Kei Kimura, Dr. Kenji Yoshino, Dr. Naoki Fujii and Dr. Takaharu Hamada of Institute of Lowland and Marine Research, Saga University for their great supports through the experiments.

References

[1] Yuji Ito, Toshiya Katano, Naoki Fujii, Masumi Koriyama, Kenji Yoshino, and Yuichi Hayami, Decreases in turbidity during neap tides initiate late winter large diatom blooms in a macrotidal embayment, Journal of Oceanography,69: 467-479. 2013.

[2] Nishikawa T (2002) Effects of temperature, salinity and irradiance on the growth of the diatom *Eucampia zodiacus* caused bleaching seaweed *Porphyra* isolated from Harima-Nada, Seto Inland Sea, Japan. Nippon Suisan Gakk 68: 356–361. (in Japanese with English abstract)

[3] Nishikawa T (2007) Occurrence of diatom blooms and damage tocultured *Porphyra* thalli by bleaching. Aquabiology 172: 405–410. (in Japanese with English abstract)

[4] Nishikawa T, Hori Y (2004) Effects of nitrogen, phosphorus and silicon on the growth of the diatom *Eucampia zodiacus* caused bleaching of seaweed *Porphyra* isolated from Harima-Nada, Seto Inland Sea, Japan. Nippon Suisan Gakk 70: 31–38. (in Japanese with English abstract)

[5] Nishikawa T, Hori Y, Nagai S, Miyahara K, Nakamura Y, Harada K, Tanda M, Manabe T, Tada K (2010) Nutrient and phytoplankton dynamics in Harima-Nada, eastern Seto Inland Sea, Japan during a 35-year period from 1973 to 2007. Estuaries Coasts 33: 417–427.

[6] Nishikawa T, Hori Y, Tanida K, Imai I (2007) Population dynamics of the harmful diatom *Eucampia zodiacus* Ehrenberg causing bleachings of *Porphyra* thalli in aquaculture in Harima- Nada, the Seto Inland Sea, Japan. Harmful algae 6: 763–773.

[7] Nishikawa T, Miyahara K, Nagai S (2000) Effects of temperature and salinity on the growth of the giant diatom *Coscinodiscus wailesii* isolated from Harima-Nada, Seto Inland Sea, Japan. Nippon Suisan Gakk 66: 993–998. (in Japanese with English abstract)

[8] Nishikawa T, Tarutani K, Yamamoto T (2009) Nitrate and phosphate uptake kinetics of the harmful diatom *Eucampia zodiacus* Ehrenberg, a causative organism in the bleaching of aquacultured *Porphyra* thalii. Harmful algae 8: 513–517.

[9] Nishikawa T, Yamaguchi M (2006) Effect of temperature on lightlimited growth of the harmful diatom *Eucampia zodiacus* Ehrenberg, a causative organism in the discoloration of *Porphyra* thalli. Harmful Algae 5: 141–147.

[10] Nishikawa T, Yamaguchi M (2008) Effect of temperature on lightlimited growth of the harmful diatom *Coscinodiscus wailesii*, a causative organism in the bleaching of aquacultured *Porphyra* thalli. Harmful Algae 7: 561–566.

[11] Syutou T, Matsubara T, Kuno K (2009) Nutrient state and nori aquaculture in Ariake Bay. Aquabiology 181: 168–170. (in Japanese with English abstract)

[12] Harada K, Hori Y, Nishikawa T, Fujiwara T (2009) Relationship between cultured *Porphyra* and nutrients in Harima-Nada, eastern part of the Seto Inland Sea. Aquabiology 181: 146–149. (in Japanese with English abstract)

[13] Arai K., T. Katano, Trend analysis of relatively large diatoms which appear in the intensive study area of the ARIAKE Sea, Japan, in winter (2011-2015) based on remote sensing satellite data, Internationa Journal of Advanced Research in Artificial Intelligence (IJARAI), 4, 7, 15-20, 2015.

[14] Arai, K., Locality of Chlorophyll-a Concentration in the Intensive Study Area of the Ariake Sea, Japan in Winter Seasons Based on Remote Sensing Satellite Data, Internationa Journal of Advanced Research in Artificial Intelligence (IJARAI), 4, 8, 18-25, 2015.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission "A" of ICSU/COSPAR since 2008. He wrote 33 books and published 500 journal papers. He is now Editor-in-Chief of IJACSA and IJISA.

Estimation of Rice Crop Quality and Harvest Amount from Helicopter Mounted NIR Camera Data and Remote Sensing Satellite Data

Kohei Arai 1
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Osamu Shigetomi [2]
Saga Prefectural Agricultural Research Institute
Saga Prefectural Government,
Japan

Masanoori Sakashita [1]
Information Science, Saga University
Saga, Japan

Yuko Miura [2]
Saga Prefectural Agricultural Research Institute
Saga Prefectural Government, Japan

***Abstract*—Estimation of rice crop quality and harvest amount in paddy fields with the different rice stump density derived from helicopter mounted NIR camera and remote sensing satellite data is made. Using the intensive study site of rice paddy fields with managing, estimation of protein content in rice crop and nitrogen content in rice leaves through regression analysis with Normalized Difference Vegetation Index: NDVI derived from camera mounted on a radio-control helicopter is made together with harvest amount of rice crops. Through experiments at rice paddy fields which is situated at Saga Prefectural Agriculture Research Institute SPRIA in Saga city, Japan, it is found that protein content in rice crops is highly correlated with NDVI which is acquired with visible and Near Infrared: NIR camera mounted on radio-control helicopter. It also is found that nitrogen content in rice leaves is correlated to NDVI as well. Protein content in rice crop is negatively proportional to rice taste. Therefore rice crop quality can be evaluated through NDVI observation of rice paddy field.**

Keywords*—Rice Crop; Rice Leaf; Nitrogen content; Protein content; NDVI

I. INTRODUCTION

Vitality monitoring of vegetation is attempted with photographic cameras [1]. Grow rate monitoring is also attempted with spectral reflectance measurements [2]. Bi-Directional Reflectance Distribution Function: BRDF is related to the grow rate for tealeaves [3]. Using such relation, sensor network system with visible and near infrared cameras is proposed [4]. It is applicable to estimate nitrogen content and fiber content in the tealeaves in concern [5]. Therefore, damage grade can be estimated with the proposed system for rice paddy fields [6]. This method is validated with Monte Carlo simulation [7]. Also Fractal model is applied to representation of shapes of tealeaves [8]. Thus the tealeaves can be asse3ssed with parameters of the fractal model. Vitality of tea trees are assessed with visible and near infrared camera data [9]. Rice paddy field monitoring with radio-control helicopter mounting visible and Near Infrared: NIR camera is proposed [10] while the method for rice quality evaluation through nitrogen content in rice leaves is also proposed [11]. The method proposed here is to evaluate rice quality through protein content in rice crop with observation of Normalized Difference Vegetation Index: NDVI which is acquired with visible and NIR camera mounted on radio-control helicopter as well as remote sensing satellite data1. Spatial resolution of Advanced Spaceborne Thermal Emission and Reflection Radiometer: ASTER/Visible to Near Infrared radiometer: VNIR is 15 m and is good enough for evaluation of rice sump density effect on quality and harvest amount.

The fact that protein content in rice crops is highly correlated with NDVI which is acquired with visible and Near Infrared: NIR camera mounted on radio-control helicopter is well reported [10]. It also is reported that nitrogen content in rice leaves is correlated to NDVI as well. Protein content in rice crop is negatively proportional to rice taste. Therefore rice crop quality can be evaluated through NDVI observation of rice paddy field. Relation among nitrogen content in rice leaves, amount of fertilizer, NDVI and protein content in rice crops has to be clarified [11]. There are some indexes which show quality of rice crops, protein content, nitrogen content, etc. in the rice leaves. Meanwhile, there are some indexes for harvest amount, the number of ear in the stump, ear length, crop weight, etc. It should be depending on circumstances of geometric condition, soil condition, meteorological condition, water supply condition, fertilizer amount and rice stump density. Intensive study paddy fields have a variety of conditions. Helicopter mounted NIR camera has a good enough spatial resolution. Therefore, rice crop quality and harvest amount is evaluated as a function of water supply condition and fertilizer amount and rice stump density. Spatial resolution of remote sensing satellite mounted visible to near infrared radiometers, however, are not so fine enough for the rice paddy fields in the sites. Therefore, only stump density influence on rice crop quality and harvest amount is evaluated with remote sensing satellite data.

[1] ASTER(Advanced Space based radiometer for Thermal Emission and Reflection)/VNIR(Visible to Near Infrared Radiometer) which is onboard Terra satellite has three spectral channels (Green, Red, and Near Infrared) with 15 m of spatial resolution.

Hiyokumochi rice is our major concern. In the intensive study paddy field, Hiyokumochi rice is planted for the research. Hiyokumochi of rice species is a new species. Therefore, appropriate grow process is unknown. One of the other purposes of this research works is to clarify an appropriate grow process, in particular, the number of ear and the number of rice in a unit area is major concern.

The method used for rice crop quality and harvest amount evaluations is described in the next section followed by experiments. The experimental results are validated in the following section followed by conclusion with some discussions.

II. PROPOSED SYSTEM

A. *Radio Controlled Helicopter Based Near Infrared Cameras Utilizing Agricultural Field Monitoring System*

The helicopter used for the proposed system is “GrassHOPPER”[2] manufactured by Information & Science Techno-Systems Co. Ltd. The major specification of the radio controlled helicopter used is shown in Table 1. Canon Powershot S100[3] (focal length=24mm） is mounted on the GrassHOPPER. The filter of blue band is replaced to the NIR filter. Therefore, Green, Red and NIR bands of images can be obtained with this camera. It allows acquire images with the following Instantaneous Field of View: IFOV at the certain altitudes, 1.1cm（Altitude=30m） 3.3cm（Altitude=100m） and 5.5cm（Altitude=150m）.

TABLE I. MAJOR SPECIFICATION OF GRASSHOPPER

Weight	2kg（Helicopter only）
Size	80cm × 80cm × 30m
Payload	600g

Radio wave controlled helicopter mounted near infrared camera imagery data is acquired at A and B paddy fields in SPRIA on 8 October 2014 with the different viewing angle from the different altitudes. Figure 5 shows an example of the acquired near infrared image. There is spectralon[4] of standard plaque as a reference of the measured reflectance in between A and B. Just before the data acquisition, some of rice crops and leaves are removed from the subsection of paddy fields for inspection of nitrogen content. Using the removed rice leaves, nitrogen content in the rice leaves is measured based on the Kjeldahl method[5] and Dumas method[6] (a kind of chemical method) with Sumigraph NC-220F[7] of instrument. The measured total nitrogen content in rice leaves and protein content in rice crops are compared to the NDVI.

B. *Rice Crop Field at Saga Prefectural Agriculture Research Institute: SPARI*

Specie of the rice crop is Hiyokumochi[8] which is one of the late growing types of rice species. Hiyokumochi is one of low amylase (and amylopectin rich) of rice species (Rice No.216).

Figure 1 shows the Location (a) and (b) Layout of the test site of rice crop field at SPRIA which is situated at 33°13'11.5" North, 130°18'39.6"East, and the elevation of 52 feet on the Google map. B and A in the Figure 1 (b) shows intensive study rice paddy fields. Figure 2 shows the superimposed image of SPARI paddy field layout on Google map. There are the test sites A and B investigation of nitrogen of chemical fertilizer dependency on rice crop quality as shown in Figure 3. There are two types of paddy subsections, densely and sparsely planted paddy fields. Hiyokumochi rice leaves are planted 15 to 20 stumps per m^2 on June 22 2014. Rice crop fields are divided into 10 different small fields depending on the amount of nutrition including nitrogen ranges from zero to 19 kg/10 a/nitrogen.

(a)Google Map

[2] http://www.ists.co.jp/?p=789
[3] http://cweb.canon.jp/camera/dcam/lineup/powershot/s100/
[4] https://en.wikipedia.org/wiki/Spectralon
[5] https://en.wikipedia.org/wiki/Kjeldahl_method
[6] https://en.wikipedia.org/wiki/Dumas_method
[7] http://www.hok-chem.co.jp/products/food/NC-220F.html
[8] https://ja.wikipedia.org/wiki/%E3%82%82%E3%81%A1%E7%B1%B3

(b)Layout of the paddy fields

Fig. 1. Location of Saga Prefectural Agriculture Research Institute: SPARI on Google map

Nitrogen of chemical fertilizer is used to put into paddy fields for five times during from June to August. Although rice crops in the 10 different small fields are same species, the way for giving chemical fertilizer are different. Namely, the small field No.1 is defined as there is no chemical fertilizer at all for the field while 9, 11, and 13 kg / 10 a/ nitrogen of after chemical fertilizer are given for No.2 to 4, respectively, no initial chemical fertilizer though. Meanwhile, 9, 11, 13 kg /10 a/nitrogen are given as after chemical fertilizer for the small field No.5, 6, and 7, respectively in addition to the 3 kg /10 a/nitrogen of initial chemical fertilizer. On the other hand, 12, 14, and 16 kg /10 a /nitrogen are given for the small fields No.5, 6, 7, respectively as after chemical fertilizer in addition to the initial chemical fertilizer of 3 kg / 10 a/ nitrogen for the small field No. 15, 17, 19, respectively. Therefore, rice crop grow rate differs each other paddy fields depending on the amount of nitrogen of chemical fertilizer.

III. Experiments

A. *Relations between Measured Indexes and NDVI Measured with Helicopter Mounted NIR Camera*

The experiment is conducted on October 3 2014 just before harvest. Figure 4 shows the NIR camera acquired image of the intensive study paddy fields A and B. There are two types of indexes, (1) rice crop quality and (2) harvest amount related indexes are measured. Those are (1) Protein Content in rice crops and SPAD (Greenness of the rice leaves) and (2) Ear Length, the Number of Ear per square meter and Crop Weight per 10a. Namely, protein rich rice crops taste bad while high SPAD means well grown rice leaves (such rice crop with high SPAD tastes good). Regression analysis is made between the measured indexes and the Normalized Difference Vegetation Index: NDVI measured with helicopter mounted NIR camera data.

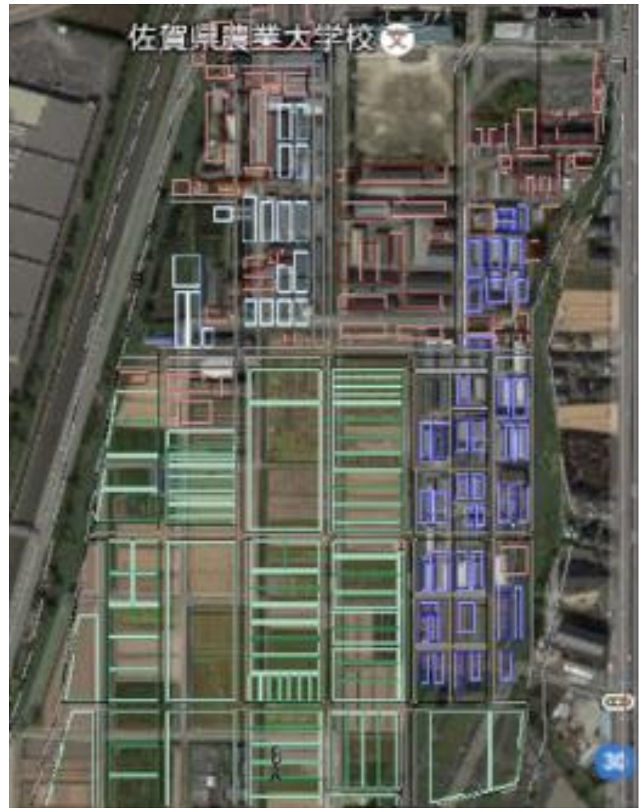


Fig. 2. Superimposed image of SPARI on Google map

A(stump density is 50/3.3m^2) B(stump density is 70/3.3 m^2)

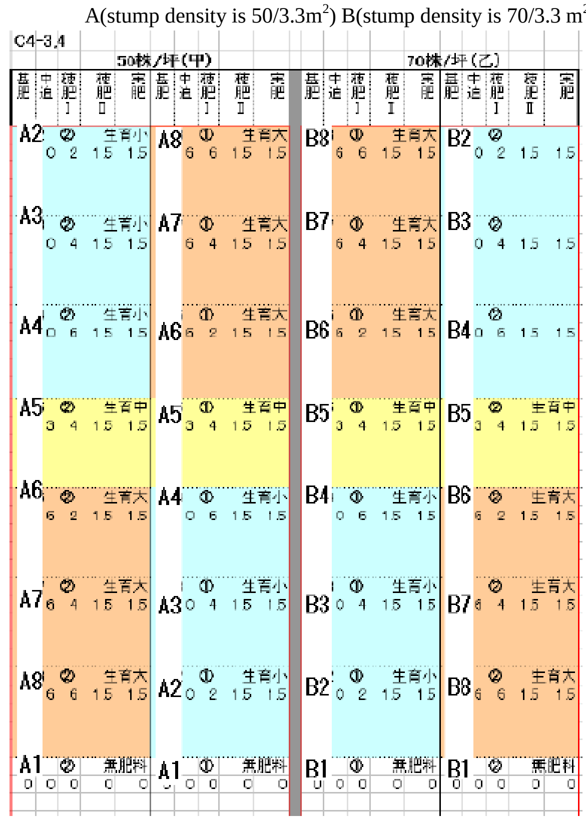


Fig. 3. Paddy filed layout for investigation of nitrogen of chemical fertilizer dependency on rice crop quality

All the indexes are measured on October 10 2014. Figure 5 shows relations between the measured NDVI and the indexes measured for the intensive study paddy field A.

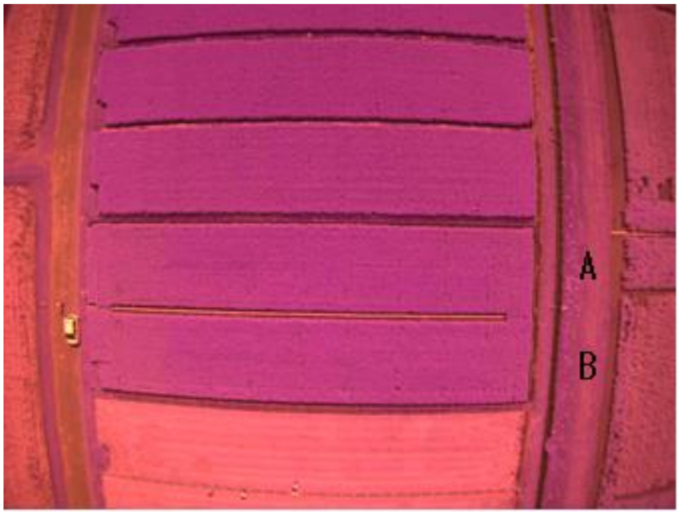
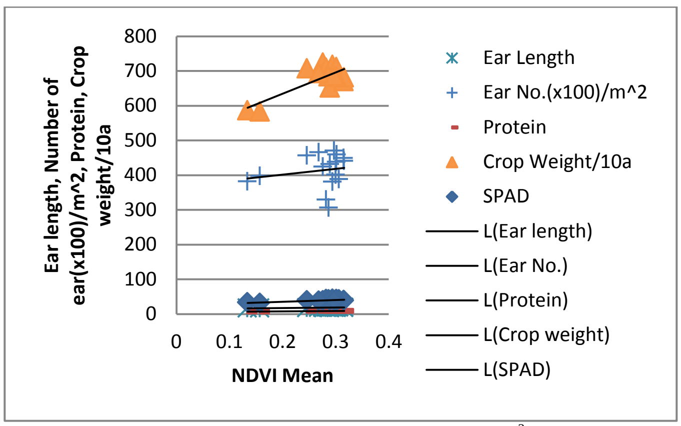
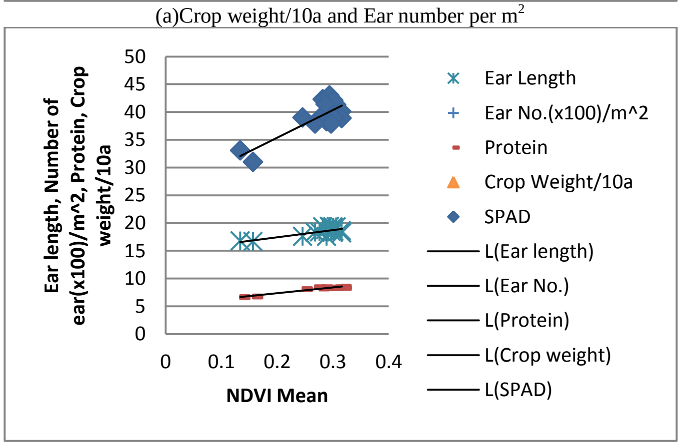


Fig. 4. NIR camera acquired image of the intensive study paddy fields A and B



(a)Crop weight/10a and Ear number per m^2



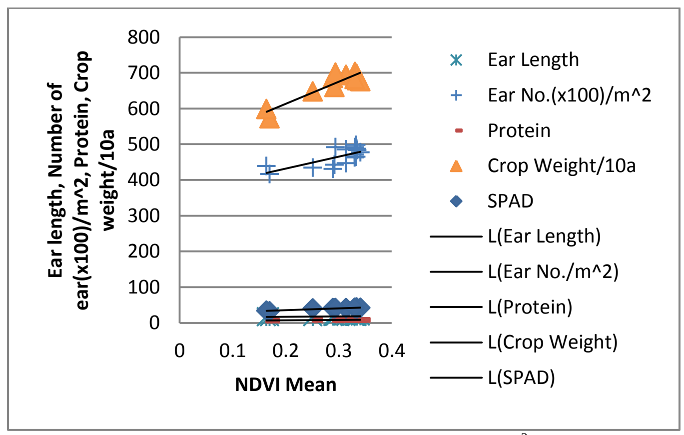
(b)Protein, Ear Length, and SPAD

Fig. 5. Rice crop quality and harvest amount related indexes

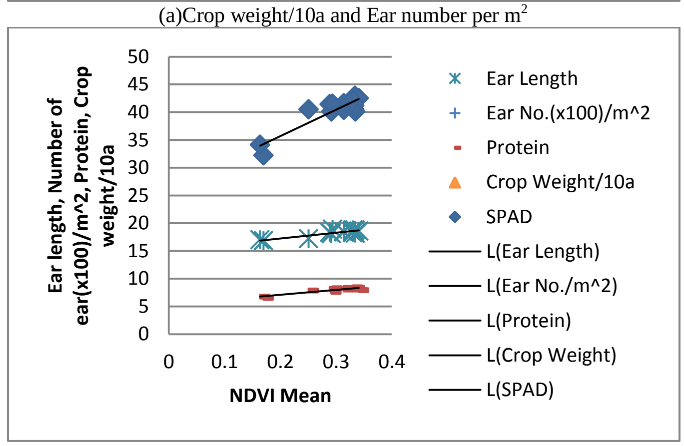
The regressive equation and R^2 for intensive study paddy field A (Stump density is $50/3.3m^2$) are shown in equation (1) to (5), respectively. Namely, the equations (1) to (5) are for the Crop Weight / 10a, the Number of Ear(x100) / m^2, SPAD, the Protein Content in the rice crops and the Ear Length.

$y = 610.6x + 512.56$
$R^2 = 0.6159$ (1)
$y = 166.44x + 368.21$
$R^2 = 0.0338$ (2)
$y = 12.99x + 14.849$
$R^2 = 0.643$ (3)
$y = 10.252x + 5.3054$
$R^2 = 0.919$ (4)
$y = 50.017x + 25.383$
$R^2 = 0.6904$ (5)

Also, Figure 6 shows relations between the measured NDVI and the indexes measured for the intensive study paddy field B. Regressive analysis is made for the intensive study paddy field B, The regressive equations and R squares shown in equation (6) to (10), respectively. In comparison of the protein content between intensive study paddy fields A and B, the protein content of A is little bit higher than that of B. Protein rich rice crops are grown in the paddy field with sparse stump density rather than dense stump density. On the other hand, crop weight / 10a of the dense stump density of paddy field is higher than the of the sparse stump density of paddy field.



(a)Crop weight/10a and Ear number per m^2



(b)Protein, Ear Length, and SPAD

Fig. 6. Rice crop quality and harvest amount related indexes

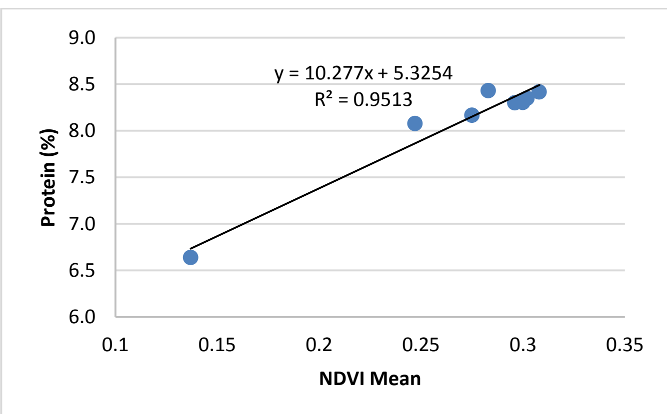
$y = 616.87x + 490.02$
$R^2 = 0.8514$ (6)
$y = 332.48x + 365.5$

$R^2 = 0.5276$ (7)
$y = 10.395x + 15.133$
$R^2 = 0.7926$ (8)
$y = 8.8038x + 5.3074$
$R^2 = 0.8134$ (9)
$y = 47.279x + 26.231$
$R^2 = 0.8168$ (10)

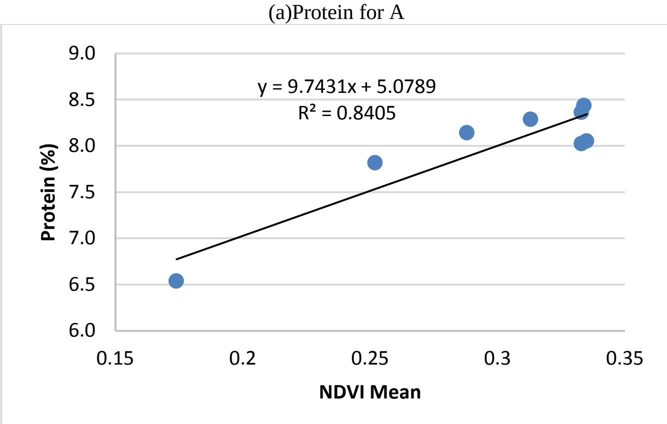
That is the same thing for the Number of Ear(x100) / m^2. Namely, the number of ear(x100) / m^2 of the dense stump density of paddy field is greater than the of the sparse stump density of paddy field while SPAD[9] of the dense stump density of paddy field is higher than the of the sparse stump density of paddy field. Meanwhile, the Ear Length of the dense stump density of paddy field is longer than the of the sparse stump density of paddy field.

B. *Relations between Protein Content and Harvest Amount and NDVI Measured with Helicopter Mounted NIR Camera*

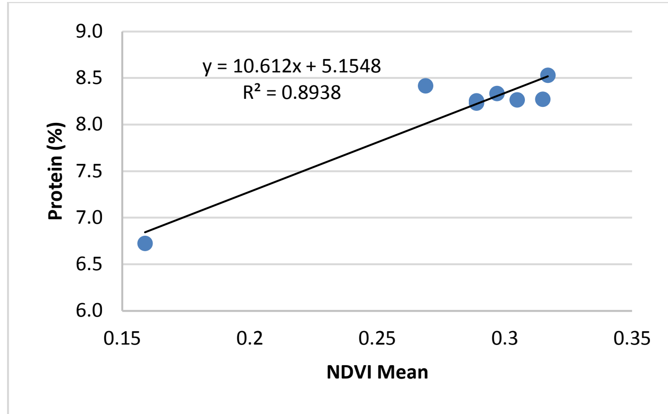
In more detail, there are 8 subdivisions in the same intensive study paddy fields A and B, A1 to A8 and B1 to B8, respectively. Also, the subdivisions are separated into two sets of subdivisions. The protein content and the harvest amount of these 16 different subdivisions are plotted in Figure 7.



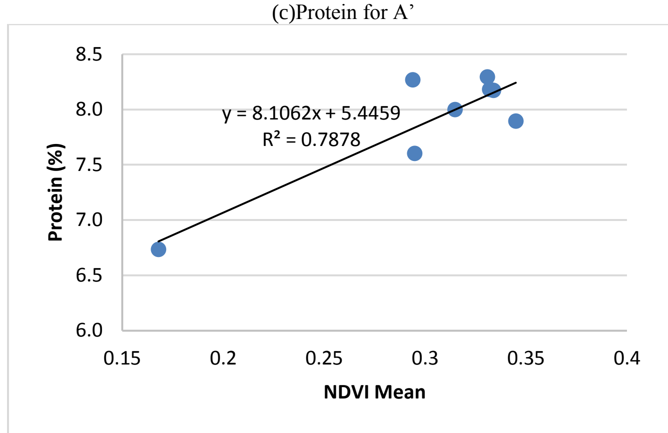
(a)Protein for A



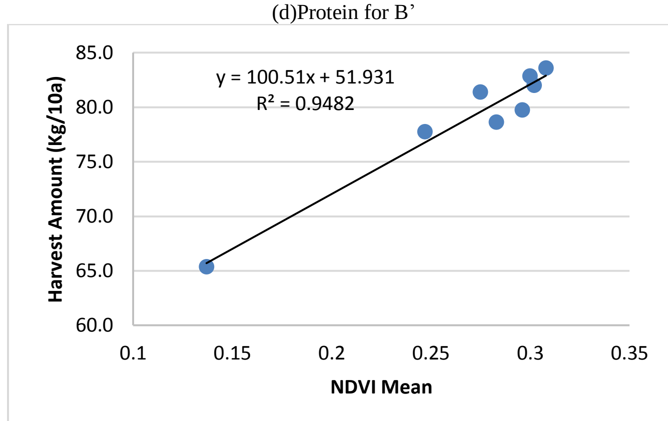
(b)Protein for B



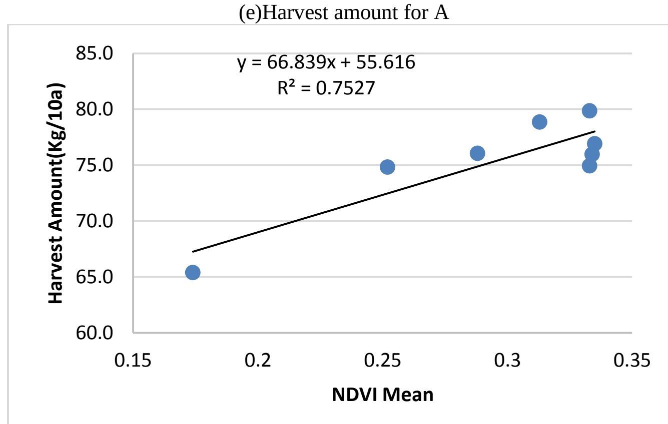
(c)Protein for A'



(d)Protein for B'

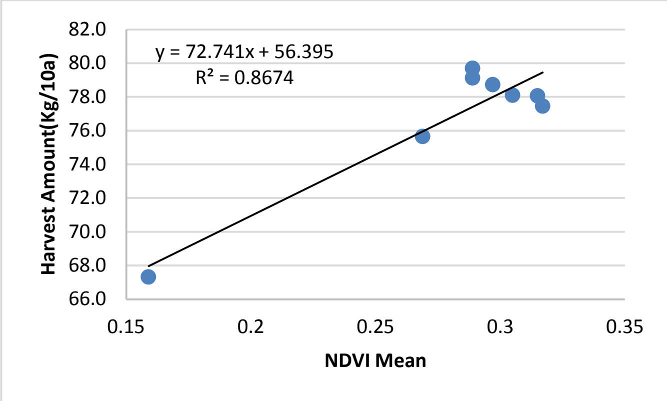


(e)Harvest amount for A

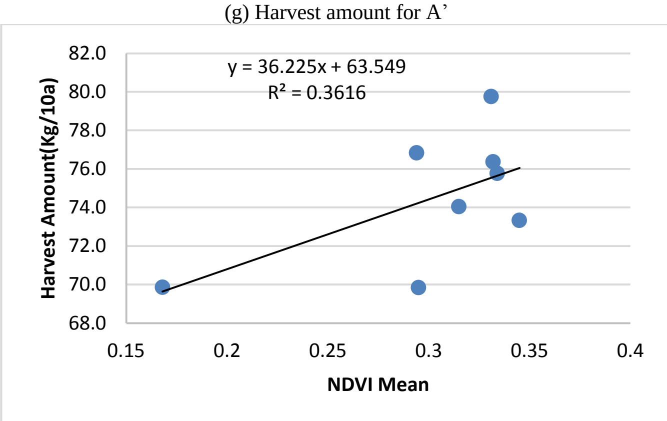


(f)Harvest amount for B

[9] Soil & Plant Analyzer Development：SPAD, chlorophyll meter



(g) Harvest amount for A'



(h)Harvest amount for B'

Fig. 7. Relations between NDVI and protein content as well as harvest amount

The relation between NDVI and protein content is better than that between NVDI and harvest amount. Also, the relations between NDVI and protein content as well as harvest amount for intensive study paddy field A (sparse stump density) are better than those for paddy field B (dense stump density).

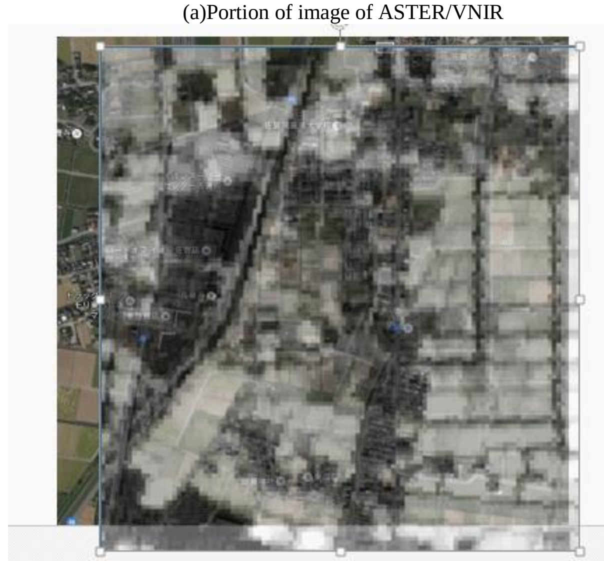
C. *Relations between Measured Indexes and NDVI Measured with ASTER/VNIR Data*

ASTER/VNIR data is acquired on October 8 2014. The small portion of ASTER/VNIR image is shown in Figure 8 (a). In the middle of the image, SPARI is situated. Figure 8 (b) and (c) shows the superimposed image of ASTER/VNIR image on Google map, the superimposed image of ASTER/VNIR image on layout map, and NDVI pixels which correspond to the intensive study paddy fields A and B, respectively. Table 2 shows NDVI of the pixels corresponding to the paddy fields A and B. NDVI of the paddy field A is 15% greater than that of the paddy field B. The result is identical to the result form regressive analysis of protein content with NDVI. These are the same for harvest amount.

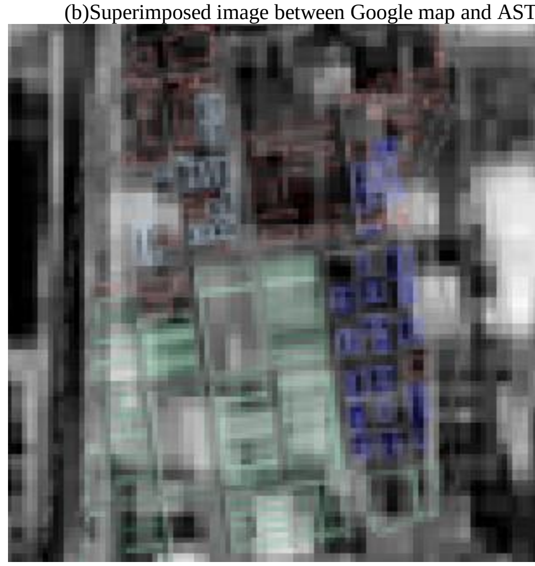
TABLE II. NDVI OF THE PIXELS CORRESPONDING TO THE PADDY FIELDS A AND B

#1	#2	#3	#4	#5	Average
0.3083	0.125	0.125	0.2286	0.2459	0.20656
0.232	0.1818	0.0952	0.0952	0.2462	0.17008

(a)Portion of image of ASTER/VNIR



(b)Superimposed image between Google map and ASTER/VNIR



(c)Superimposed image between layout map and ASTER/VNIR

Fig. 8. Images of ASTER/VNIR and Google map and layout image as well as pixels corresponding to the paddy fields A and B

Therefore, the relation between NDVI and rice crop quality which is represented with protein content in rice crops is confirmed with ASTER/VNIR remote sensing satellite data.

D. *Most Appropriate Number of Ear and Number of Rice /m²*

Relation between NDVI and one of the other concerns of the number of rice / m^2 is investigated. Figure 9 shows the relation. Although R^2 is not high enough, it is confirmed that the number of rice / m^2 is proportional to NDVI. Therefore, it is found that around 322 and 348 of the number of rice / m^2 would be appropriate for stump density 50 and 70 / 3.3 m^2, respectively. Also, as shown in Figure 7, it is found that 414 and 464 of the number of ear(x100) / m^2 are appropriate for stump density of 50 and 70 / 3.3 m^2, respectively. These are estimated with the measured NDVI.

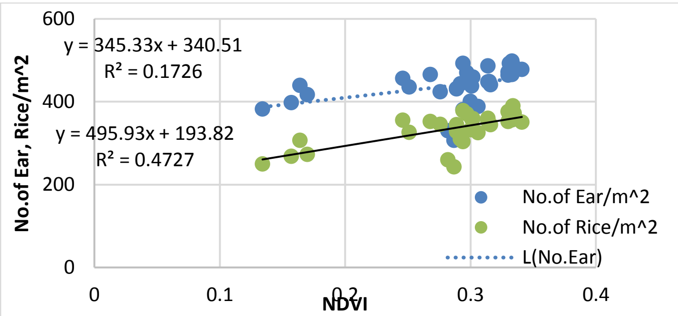


Fig. 9. Relation between NDVI and one of the other concerns of the number of ear and rice / m^2

IV. CONCLUSION

Estimation of protein content in rice crop and nitrogen content in rice leaves through regression analysis with Normalized Difference Vegetation Index: NDVI derived from camera mounted radio-control helicopter and remote sensing satellite mounted visible to near infrared radiometers made. Through experiments at rice paddy fields which is situated at Saga Prefectural Agriculture Research Institute: SPARI in Saga city, Japan, it is found that protein content and harvest amount of rice crops is highly correlated with NDVI which is acquired with visible and Near Infrared: NIR camera mounted on radio-control helicopter and visible to near infrared radiometer onboard remote sensing satellite.

Protein content in rice crop is negatively proportional to rice taste. Therefore, rice crop quality can be evaluated through NDVI observation of rice paddy field. It is found that harvest amount is linearly proportional to NDVI. Therefore, it is possible to estimate rice crop quality and harvest amount with NDVI. It is also found that rice crop quality of low density of rice stump is better than that of highly dense rice paddy field. It is found that around 322 and 348 of the number of rice /m^2 would be appropriate for stump density 50 and 70 / 3.3 m^2, respectively. Also, it is found that 414 and 464 of the number of ear(x100) / m^2 are appropriate for stump density of 50 and 70 / 3.3 m^2, respectively. These can be estimated with the measured NDVI.

ACKNOWLEDGEMENT

Authors would like to thank Dr. Satoshi Tsuchida and the other related members of National Institute of Advanced Industrial Science and Technology: AIST, Japan for their providing ASTER/VNIR.

REFERENCES

[1] Wiegand, C., Shibayama , M, Yamagata , Y, Akiyama, T., 1989. Spectral Observations for Estimating the Growth and Yield of Rice, Journal of Crop Science, 58, 4, 673-683, 1989.

[2] Kohei Arai, Method for estimation of grow index of tealeaves based on Bi-Directional reflectance function:BRDF measurements with ground based netwrok cameras, International Journal of Applied Science, 2, 2, 52-62, 2011.

[3] Kohei Arai, Wireless sensor network for tea estate monitoring in complementally usage with Earth observation satellite imagery data based on Geographic Information System(GIS), International Journal of Ubiquitous Computing, 1, 2, 12-21, 2011.

[4] Kohei Arai, Method for estimation of total nitrogen and fiber contents in tealeaves with grond based network cameras, International Journal of Applied Science, 2, 2, 21-30, 2011.

[5] Kohei Arai, Method for estimation of damage grade and damaged paddy field areas sue to salt containing sea breeze with typhoon using remote sensing imagery data, International Journal of Applied Science,2,3,84-92, 2011.

[6] Kohei Arai, Monte Carlo ray tracing simulation for bi-directional reflectance distribution function and grow index of tealeaves estimation, International Journal of Research and Reviews on Computer Science, 2, 6, 1313-1318, 2011.

[7] Kohei.Arai, Fractal model based tea tree and tealeaves model for estimation of well opened tealeaf ratio which is useful to determine tealeaf harvesting timing, International Journal of Research and Review on Computer Science, 3, 3, 1628-1632, 2012.

[8] Kohei.Arai, H.Miyazaki, M.Akaishi, Determination of harvesting timing of tealeaves with visible and near infrared cameradata and its application to tea tree vitality assessment, Journal of Japanese Society of Photogrammetry and Remote Sensing, 51, 1, 38-45, 2012

[9] Kohei Arai, Osamu Shigetomi, Yuko Miura, Hideaki Munemoto, Rice crop field monitoring system with radio controlled helicopter based near infrared cameras through nitrogen content estimation and its distribution monitoring, International Journal of Advanced Research in Artificial Intelligence, 2, 3, 26-37, 2013.

[10] Kohei Arai, Masanori Sakashita, Osamu Shigetomi, Yuko Miura, Estimation of Protein Content in Rice Crop and Nitrogen Content in Rice Leaves Through Regression Analysis with NDVI Derived from Camera Mounted Radio-Control Helicopter, International Journal of Advanced Research in Artificial Intelligence, 3, 3, 13-19, 2013

[11] Kohei Arai, Masanori Sakashita, Osamu Shigetomi, Yuko Miura, Estimation of Protein Content in Rice Crop and Nitrogen Content in Rice Leaves Through Regression Analysis with NDVI Derived from Camera Mounted Radio-Control Helicopter, International Journal of Advanced Research in Artificial Intelligence, 3, 3, 13-19, 2013

AUTHORS PROFILE

Kohei Arai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He is the Edito-in-Chief of IJACSA and IJISA. He wrote 33 books and published 510 journal papers as well as 450 conference papers.

http://teagis.ip.is.saga-u.ac.jp/

A Design of a Multi-Agent Smart E-Examiner

Khaled Nasser ElSayed
Computer Science Department, Umm AlQura University

***Abstract*—this paper proposes a design of an application of multi agent technology on a semantic net knowledge base, to build a smart e-examiner system. This e-examiner could be used in building and grading a personalized special on-line e-assessment. The produced e-assessment should cover the majority of examined topics and material. It should cover various levels of difficulties and learners profile(s). The e-examiner will use a semantic net question bank, to emphasize on the structuring categories of all course domains. This task is done through four different intelligent agents: control agent, personal agent, examiner agent, and grading agent. The system might select questions from a bank of questions for several courses. It could be used in different education levels and natures. Also, it will produce a key for the produced exam, to be used latter in grading, and giving final marks of e-assessments.**

Keywords—m-Learning; e-Assessments; Multi-agent; Semantic net; Examiner

I. Introduction

The e-learning and m-learning are considered the important ways to respond to the needs of the d-learning. Many traditional learning styles should be applied together with the learner-centered approach to represent the needs of the e-learning system [1].

Software educational systems have become highly complex in both concept and design in the education of most sciences. In these systems, there is a wide difference between the role of the computer from solely a transmitter of knowledge and using it as a tool that aids in the construction of knowledge, as in [2].

There are an increasing number of AI algorithms and applications that are applied in tutoring and education systems. One of the AI techniques is the intelligent agent, which could be widely used in education. An intelligent agent is an autonomous calculated entity.

An intelligent agent could collect personal interests of the learner (s) and user (s) to give instance response according to the pre-specified demands of them. A personal one can discover their personal interests easily without disturbing them. So, it is suitable for personalized e-learning by recommending selected suitable learning materials [3].

An agent could detect by its sensors and perform actions on its environment by actuators to achieve target goals. The intelligent agent could learn new knowledge and uses it in achieving target goals. It might be simple procedure or program to perform simple goal or task. Also, it could be a complicated system working as a group of persons co-operated together towards a target goal [4].

A multi-agent system is constructed from a several intelligent agents, cooperated and interacted together to achieve target goal(s) within an environment. It could solve problems that are difficult be solved with an individual agent [5].

Adaptive learning is a learner-centric and individualized learning. The target of this learning may be certain learner or group of learner have a public feature(s). According to that, each learner could select chapters, exercises, and remedial actions even the curriculum relative to his individual's mental state. Then the learning system can maintain a dynamic knowledge for learners. It monitors the learner's behavior and stores his mental model which could be used in identifying the causes of the student performance [6].

Nowadays, there are many tools help in producing assessments. Advanced eLearning Builder [7] is an Authoring tool designed for creating e-learning materials such as e-tests, tutorials and quizzes in visual mode and compiling them in EXE form. Also, MicroGrade and Micro Test [8] are desktop applications with built-in access to hosted Internet sites for secure test delivery and grade posting for Windows and Macintosh compatible. While, OnlineTesting offers in [9] software and programming services to offer online quizzes and tests through the Internet or within an Intranet.

There are many other software tools for producing assessments like; ProPrfos [10], Adit Sofware [11], Quizworks [12] and ExamJet [13]. Majority of those software tools enable the human examiner to select questions or the question are selected randomly.

This paper presented a design of a multi-agent based smart e-examiner system. This design is for building a personalized e-assessment according to the study, level and profile of the learner(s). It builds its semantic net as object oriented knowledge base (question KB) to be filled with questions that cover the materials of topics of courses. Then it is running its four agents (control agent, personal agent, examiner agent and grading agent) on that question KB to select questions and building e-assessments according to learner(s) education level and profile data. The e-examiner agent gets its intelligence from watching and learning from the human examiner to enhance the selection of questions. It is designed for education in academic institutions, training centers, human resources departments. Also, it can help human examiners in producing their own assessments.

II. Categories of Generated Questions

Questions are the infrastructure of the system, which could be generated for building an e-assessment. These questions and their answers from several e-courses construct the semantic net question KB. There are four categories of exercises; each includes a number of questions of that category.

- **The Multiple Choice Category**; where a question could be a text (sentence), while the answer is one correct choice from available four multiple choices (might be less). The answer of this category will be checked and graded by the e-examiner system, as shown in Figure 1-a.

a.

1	**Which of the following languages have case-sensitive names?**
T	❑ C++
	❑ COBOL
	❑ Fortran
	❑ Pascal

b. Verifies whether or not the string can be generated by grammar.

Answer: ☐ **True** ☐ **False**

c. A………. determines the order in which operators of equal Precedence are evaluated when they occur in the same expression

(Associativity, Availability, Precedence, Mandatory)

d. Consider the following XX **grammar:**

LP → LP FUNC (PAR ; PAR) | LP (PAR) | exec
FUNC → ID
PAR → LP | NUM

Note: (**NUM** is the token for numbers and **ID** is the token for identifiers).

Answer the following questions:

- Give the equivalent grammar after removing the left recursion from the above grammar
- Derive the following string from LP.

exec svg (exec max (2 ; 5) ; 4)

Fig. 1. Examples of Question Types stored in Question KB

- **The True/False Category**; where a question could be a text (sentence), while the answer should be one of two options (true or false), it is a logical choice. The answer of this category will be checked and graded by the e-examiner system. An example of this category is shown in Figure 1-b.
- **The Fill in the Blanks Category**; where a question could be a text (sentence), while the answer is one-word/simple-sentence to be selected from a pool-from 2 to 10- of words/small sentences. So, the process of filling in the spaces looks like matching certain question with a suitable answer. The answer of this category will be checked and graded by the e-examiner system. An example of this category is shown in Figure 1-c.
- **The General Category**; where a question could be in a form differs than each of the above 3 forms. It could be a long or short text, including graphs, equations, or numbers. Also, it could be imported from an external file. While the answer could be a free text including graphs, equation, or whatever the examined person want to write. Answer of this category of question should be checked and graded manually, by the examiner/ instructor. An example of this category is shown in Figure 1-d.

III. MAIN COMPONENTS OF THE E-EXAMINER

The proposed e-examiner system consists mainly, as shown in Figure 2, of four components; each component cooperates with each other to prove the system generality.

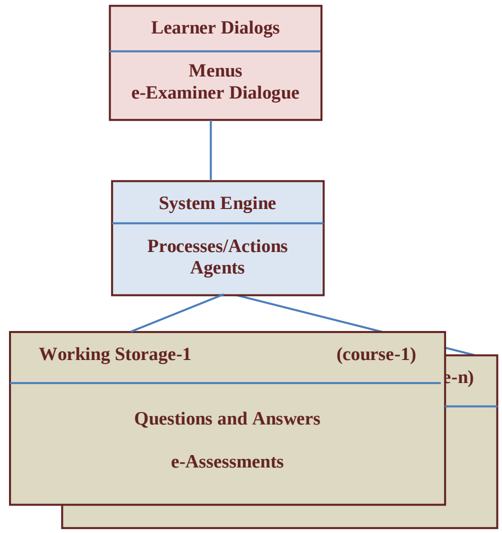


Fig. 2. Structure of the e-examiner system

- **Learner Dialog**; which is a menu-driven dialog for easing learner interaction with the e-examiner system directly. After finishing certain topics or certain course or part of course, he will navigate this dialog, thin the e-examiner will start creation of a personalized e-assessment special to that learner. After passing answers of a learner to the e-examiner, it will check and grade his answers, and update the profile of that learner.
- **System Engine**; which involves four intelligent agents, as shown in Figure 3. The first agent is the control agent, which controls the process and agents. The second intelligent agent creates, fills and updates personal profiles of the examined learners. While the third agent will serve in selecting questions and create a personalized e-assessment and learns from the instructor examiner by watching him and gets from his experience, at editing questions or at creating exams. The fourth agent is the grading agent, which is responsible of checking and grading learner answers.
- **Question Bank Knowledge Base**; which is a semantic network KB, as shown in Figure 4. It consists of nodes like: exams, exercise, questions, answers and links (relations between nodes). Its structure is based on object oriented analysis of application domains. It holds an indexing structure used by the inference engine to locate questions. Each node or link is represented by an object of a suitable class.
- **Working Storages**; where each working storage is assigned to certain topic, chapter or course. The e-examiner system can be switched from one working storage to another.

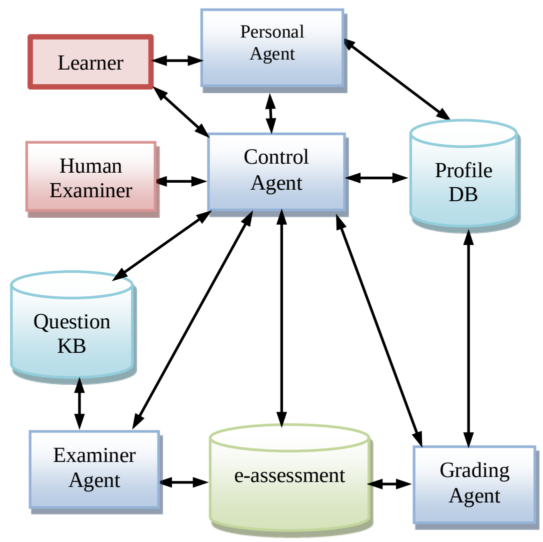


Fig. 3. Intelligent Agents in e-examiner system Agent

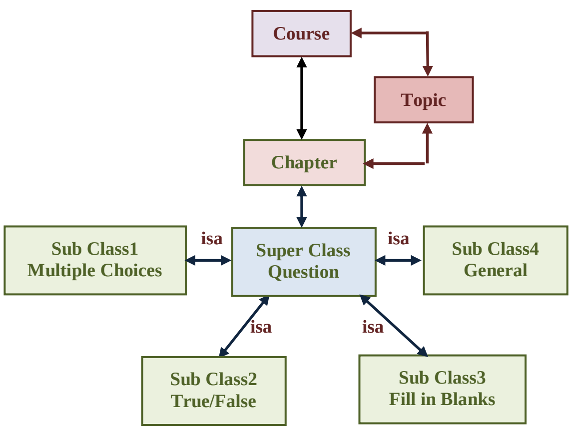


Fig. 4. Semantic Net of e-Examiner System

IV. Building the Question KB

A topic (course) material as application domain is analyzed using object oriented techniques to be used to create e-assessment. This e-assessment should cover that material completely or even a part(s) of it the learner profile and studying. Each domain is defined as objects like courses, topics, chapters, assessment exercises, questions, attributes, answers and relationships between these objects.

This question KB is designed, as shown in Figure 4, to be instantiated for several course, one course at a time. In the first use of the e-examiner, name of college or institute, department, and courses could be identified. Then, questions and their answered could be edited.

A. Organization of the Question KB

Questions to select from, should be distributed over the question KB, which designed and organized in certain manner that enables the e-examiner to select questions randomly according the learner(s) profile and e-assessment factors, which are supervised by the intelligent agents.

At editing questions to be acquired, the editor will provide the KB with certain information to be used later as a basis for the e-examiner to select a set of questions for certain e-assessment. It should takes in consideration, that the e-assessment should cover majority or even specified parts of course material and should include various levels of difficulties.

The KB is designed to deal with the categories of question listed in the previous section. Its entity classes are fulfilled with attributes that lead it to be suitable to acquire the knowledge learned from the examiner at building e-assessments. It has one Super Class and 4 Sub Classes, which has to be suitable to hold all categories of question listed in the previous section.

*1) **The Super Class question** is the base for all categories of questions. It has some important attributes that are inherited by the other four sub classes. These attributes include a string attribute for the question text, pointer to the chapter or topic related to, and pointers to two lists of links (relations). It also, includes an attribute for learner answer of that question.*

The first type of links are to the reminded questions – questions that are suggested to be with that a certain question in the same e-assessment. While, the second type of links are to the rejected questions –questions that are not allowed to be with that a certain question in the same e-assessment. The other four sub classes should have the inherited attributes of the super class, in addition to special attributes as listed below.

*2) **The Sub Class for Multiple Choice Questions**; which have many special attributes: four string attributes for predicted choices, and character attribute to denote the correct choice.*

*3) **The Sub Class for True/False Questions**; which have only one special attribute, which is a binary attribute. This attribute will hold the answer, which is 1 (for True question text) or 0 (for False Question text).*

*4) **The Sub Class for Filling in Blanks Questions**; which have only one special attribute, which is a string attribute to hold the correct text to be filled in the spaces.*

*5) **The Sub-Class for General Questions**; which have only one special attribute, to hold the complete answer of the generic question. Contents of this attributes could be imported from an external file. This answer could be edited or imported from a file.*

B. Filling the Question KB

Question KB should be initiated for each course individually by certain information, like course code and name, department and/or institute offering that course; number of predicted chapter (could be modified latter by increasing or decreasing certain chapter). For each chapter, questions and their answers according to the four previous question types, listed before in section II, could be supplied to the system.

The human examiner starts supplying questions to the e-examiner. First, he selects the course, topic, chapter, and one question type of those listed in section II. Then, he supplies the question text, accompanied with all needed information as described in section III. Then, e-examiner will fill the KB tables with questions and their related knowledge.

Knowledge within the two lists of links could be acquired at entering a new question or be learnt from the human examiner when he add/rejects certain question.

V. Intelligent Agents and an E-Assessment

The process of building an e-assessment should emphasize that each course include a number of chapters/topics. Some or all of these chapters could be covered by the e-assessment. This will help the e-examiner in selecting questions from the specified chapters or topics. Also, dividing a course into chapters helps to divide assessment into logical parts and "accumulate" assessment question KB (for example, you have 500 questions, but single e-assessment consists of 20 questions that are randomly selected from the KB).

A. Role of Agents in the e-Examiner System

As seen in Figure 3, the e-examiner system has four agents; each of them has its important role in the main process of building and grading of an e-assessment, as described below:

- **Control Agent (CA**) is responsible of managing the process of building and grading of e-assessments. It controls and initiates other agents.
- **Personal Agent (PA**) is responsible of gathering information about learner and saving it in profile file. It also retrieves and updates his profile file.
- **Examiner Agent (EA)** is responsible of **selection** of question from question KB. It uses random and mathematical functions in additions to the knowledge resides in the relations between different questions to pick the best choices from many enumerated possibilities. It manipulates questions and other objects and the in-between relations in a standard way with all presented topic(s) or chapter(s). It watches the examiner when he requests to remove/add certain questions from/to the e-assessment produced by the system (in the training phase). It has its actions on the knowledge in the relation between questions for next e-assessment building.
- **Grading Agent (GA)** is responsible of checking and grading e-assessments passed by certain learner(s). It compares answers of learner with answers stored with questions in question KB. Then, it updates profiles of examined learner(s), according to grading results.

B. Criteria of an e-Assessment

The presented e-examiner is a comprehensive solution for building e-assessments on the Internet and Intranet and grading those e-assessments. It can easily select all exercises types and format the text, add graphics, formulas in offline and online e-assessments.

In the offline e-assessments, the tool can produces an e-assessment to be printed and copied to be distributed over students in a class. Also, it will produce a key for each e-assessment, to help the examiner to check answers of the examined people, and grading their e-assessment.

E-examiner does the process of building an e-assessment, as shown in Fig. 5, it emphasized the following criteria:

- The heading interface of the e-assessment is fully customizable.
- E-examiner supports multi-language.
- Questions and answers could be shuffled randomly (for every particular question), to avoid any fraud by people who pass your tests.
- Unlimited number of questions in a single test and up to 4 answers in multiple choices.
- E-examiner system is trained in the beginning under supervision of the human examiner.
- Then the whole process is fully automated.

C. Initiating the e-examiner System

Before using the e-examiner system on-line with learners to generate e-assessment(s), it has to be initialized and trained to be ready for its work. Also, it will be adjusted to produce off line assessments.

For producing assessments using e-examiner system, the human examiner should initiate the assessment by supplying some information like the name, top introduction, bottom conclusion, and logo of the examining institute. Then, he should provide course name and chapters to be covered in the assessment to be selected from those available in the question KB.

Some optional data -could be applied to the system like: Academic year, Semester, Student sections, e-assessment duration, maximum score, e-assessment type (midterm, quiz, final), sort of e-assessment (A, B), total mark. While for online/offline assessment, the human examiner should specify the types of exercises in the assessment, number of question in each, and the mark specified for each exercise.

VI. Building and Grading an E-Assessment

The e-examiner system can start production of offline/online assessments under the supervision the human examiner. Then it can generate e-assessment alone with no need to guidance from human examiner.

A. Training of the e-Examiner by Human Examiner

In first use of the e-examiner system, it should passes through a training session to build e-assessments under the supervision of a human examiner, as seen in the algorithm shown in Figure 5.

The EA agent will get its experience by watching the human examiners, in two different situations.

First, at editing a question; when the human examiner specifies that some question(s) are preferred to be in the same e-assessment or should prevented from being in the same e-assessment with that question.

Secondly, at building an e-assessment, when the human examiner asks to remove certain question(s) or adds certain ones.

1. **PA agent GETS e-Assessment Data:** Course, Chapters covered, From Learner(s) profile
2. **CA agent GETS Exercise Data:** number of exercises, type, number of questions in each exercise from Human examiner
3. **CA agent CALCULATES** Number of questions in each exercise from each Chapter.
4. **For each exercise in the e-assessment Do**
5. **Do Until end of questions in the exercise**
   EA agent SELECTS a question using Random function & Links
   EA agent CHECKS Reminding & Rejection links in the relationships of the question KB net.
6. **CA agent DISPLAY** Complete exercise to the human examiner.
7. **IF NOT APPROVED** (by human examiner)
8. **EA agent GETS** rejections and reasons, suggestions and reasons (reminding), **UPDATE** links of the question KB, and **UPDATES** exercise.
9. **ELSE CONTINUE** (go to 4)
10. **CA agent CONSULTS complete e-Assessment to human examiner**

\

Fig. 5. Algorithm for Training the e-Examiner System

At beginning a session to build an e-assessment from the question KB, e-examiner system performs its task through several steps, which incorporates knowledge acquisition and learning with knowledge retrieval from its question KB, based on intelligent agents.

The e-examiner asks the examiner about type of each exercise and its number of questions. It calculates number of questions per each chapter by mathematical functions.

Then, EA agent starts a process of selection questions, one by one, until finishing selecting. This selection is based on random function and according to the reminding and rejecting knowledge resides in the links between each question and other questions in the question KB.

The e-examiner consults selection results in a complete exercise to the human examiner for approval. Sometimes, she/he removes certain question or adds another question. In this case, the EA agent asks her/him for explanation.

This explanation is acquired to a rejecting knowledge in case of refusing certain question, and to a reminding knowledge in case of adding certain questions. This knowledge will reside in the links between any requested question and other questions in the question KB.

After performing the requested modification, the tool, finally, produces a file including the final form of the e-assessment or the quiz. This e-assessment version should be reviewed by the human examiner. Also, the e-examiner system produces a key for questions.

B. *The e-Examiner Builds an e-Assessment*

After the e-examiner passed about 50 training sessions in building e-assessments, under the supervision of a human examiner, it is ready to build its own e-assessment, depending on his knowledge. It builds its e-assessments, as shown in Figure 6, depending on the learner profile data and the regulations acquired from before the human examiner.

C. *e-Examiner Grades an e-Assessment*

The grading process of an e-assessment is done so easily by the e-examiner system. At selecting certain question to add to an e-assessment, the EA agent retrieves its answer from the question KB, attached with it. Then the GA agent compares the answer of learner with the answer retrieved from the question KB. If they are the same, it accumulates a one degree to the total degrees collected by the learner in that e-assessment. Finally, it gives the total degree to the PA agent to update the learner profile data.

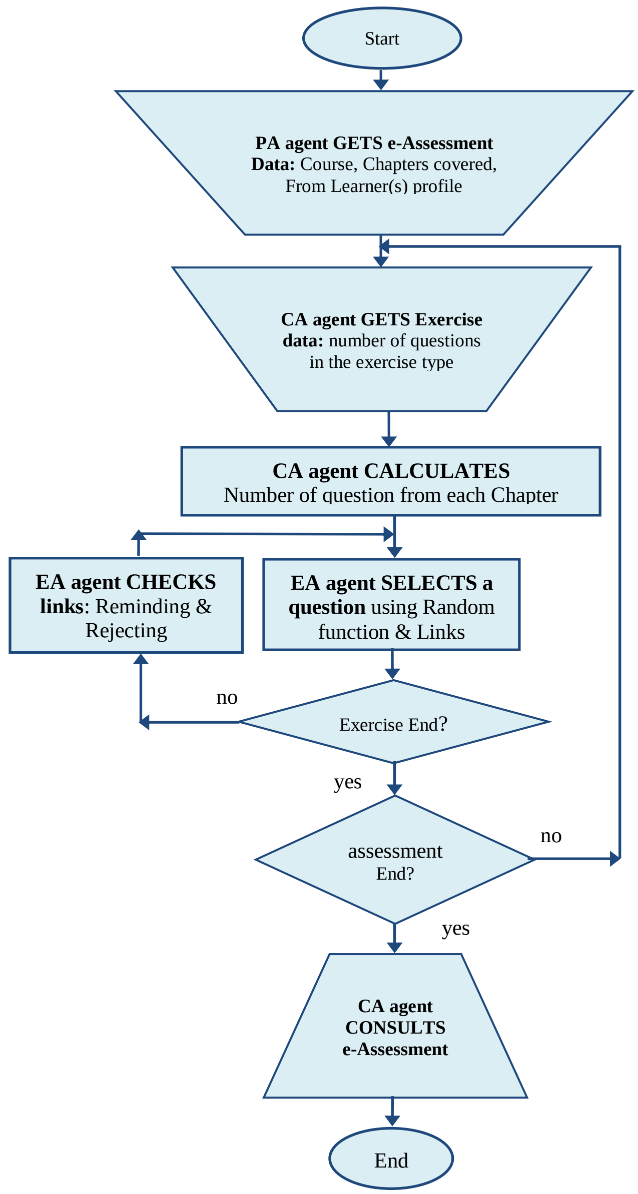


Fig. 6. Building an e-assessment by the e-examiner system

VII. CONCLUSIONS

The presented paper presented a design of a smart e-examiner system, which is running four intelligent agents on a semantic net of question KB. Its agents; control agent, personal agent, examiner agent, and grading agent co-operated together to perform its task of producing and grading personalized e-assessment, depending on learner studying and level.

The system should start its work under the supervision of human examiner. The examiner agent (EA) got experience by watching the human examiner, in two different situations: at editing a question; when the he specifies that some question(s) are preferred to/not-to be prevented from being in the same e-assessment with that question, and at building e-assessment; when the examiner asks to remove certain question(s) or add certain ones.

The e-examiner system could be used on various education level and nature to generate e-assessments for employees and students in training courses. Also, it produced a key for the produced e-assessment, to be used in grading, and giving final marks of e-assessments. Future work will include enhancing system architecture and improving its security and reading IP address of learner.

REFERENCES

[1] Zhi Liu & Bo Chen, "Model and Implement an Agent Oriented E-Learning System", the International Conference on Computational Intelligence for Modeling, Control and Automation, and International Conference Intelligent Agents, Web Technologies and Internet Commerce, Vienna (CIMCA-IAWTIC'05) 0-7695-2504-0/05, 28-30 Nov.2005, pp. 859–864. http://doi.ieeecomputersociety.org/10.1109/CIMCA.2005.1631576

[2] J. Anderson, F. Boyle, A. Corbett, and M. Lewis, "Cognitive Modeling and Intelligent Tutoring", Elsevier Science Publishers, Journal Artificial Intelligence, Special issue on artificial intelligence and learning environments archive, vol. 42, no. 1, pp. 7-49, Feb. 1990. http://act-r.psy.cmu.edu/wordpress/wp-content/uploads/2012/12/119CogMod_IntTut.pdf

[3] Jin-Ling Lin and Ming-Hung Chen, "An Intelligent Agent for Personalized E-Learning", 8th Inter. Conference on Intelligent Systems Design and Applications - Volume 01, Kaohsiung, 26-28 Nov. 2008, pp. 27-31. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4696172

[4] S. J. Russell, and P. Norvig, (2009, Dec. 11), *Artificial Intelligence – A Modern Approach*, 3rd Ed., Prentice Hall. http://stpk.cs.rtu.lv/sites/all/files/stpk/materiali/MI/Artificial%20Intelligence%20A%20Modern%20Approach.pdf

[5] M. Niazi and A. Hussain, "Agent-based Computing from Multi-agent Systems to Agent-Based Models: A Visual Survey". Springer Scientometrics, vol. 89, no. 2, pp. 479–499, 2011. doi: 10.1007/s11192-011-0468-9.

[6] S. A. Agarwal and M. Sinha," Multi Agent Adaptive E-Learning System", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 7, pp. 242-246, IJARCSSE, July 2013.

[7] Advanced eLearning Builder, "Authoring tool designed for creating e-learning materials", 2015. http://www.eduiq.com/elearning.htm

[8] MicroGrade and MicroTest,"desktop applications", 2015. http://www.chariot.com/

[9] OnlineTesting,"software and programming services", 2015. http://www.onlinetesting.net/

[10] ProfProf, "Create Online Tests & Quizzes Easily", 2015. http://www.proprofs.com/quiz-school/

[11] AditSofware," Exams, Tests and Quizzes Made Easy", 2015. http://www.aditsoftware.com/

[12] Quizworks, " Build online exams with our fun & easy to use tool", 2015. "https://www.onlineexambuilder.com/

[13] ExamJet Quiz Maker," Quiz maker software made with love", 2015. https://www.igneon.com/

Improved Text Reading System for Digital Open Universities

Mahamadou ISSOUFOU TIADO
Department of Mathematics and Computer Science,
University of Abdou Moumouni, UAM
Niamey, Niger

Abdou IDRISSA
University Institute of Technology
University of Tahoua
Tahoua, Niger

Karimou DJIBO
Department of Mathematics and Computer Science,
University of Abdou Moumouni, UAM
Niamey, Niger

***Abstract*—The New Generation of Digital Open Universities (DOUNG) is a recently proposed model using m-learning and cloud computing option and based on an integrated architecture built with open networks as GSM and Internet. The goal of achieving the ubiquitous ability of the m-learning is having the large number of languages as a serious issue. It needs to use many teachers in order to repeat the same course in various languages. In this paper, an extended system is proposed under the consideration of the low capacities of the cell-phone device in terms of computing and visualization. The model uses the possibility to build a voice warehouse which can be used to generate the audio format of every course provide in a text format and in a particular language. The Advanced Text Reading System (ATRS) is proposed to use that voice warehouse and to produce the audio format of a course, giving facility to teachers to easily overcome the constraints of language barrier. The new proposed model is described and its contributions are discussed.**

***Keywords*—m-learning; distance learning; digital open universities; cloud-computing; audio warehouse**

I. Introduction

The advent of the New Generation of Digital Open Universities was previously proposed to achieve the goal of realizing the conversion of the traditional distance learning to digital solutions with multiple options. Some of the DOUNG options allow to record the sessions given by a teacher in a multimedia format or to convey the multimedia stream instantly through the network towards the learners. In this operating model, the saved multimedia files contain a particular course in a specific language. An incoming limitation is related to the need to deliver the same course in another language. The nowadays solution uses another session given by a teacher chosen according to the targeted language, with also another work of recording to do. That operating model becomes heavier and generates more issues when addressing large public of learners around the world according to the ubiquitous nature of the m-learning. Thus, the learner's language becomes a barrier for the teacher and another parameter for recording new sessions. The goal of providing courses to a wide public of learners around the world in as many languages as possible brings to the proposition of the audio warehouse and its exploitation script. They are used to facilitate the conversion of the courses in audio format within a selected language. In addition, this new operating model that generates audio format of courses meets the constraints of the extension of the DOUNG services over GSM network.

After presenting the basic concepts of the distance education in point II, the hybrid model of the DOUNG is presented in point III. The point IV presents the voice warehouse with its building process. The point V highlights the ATRS algorithm used on the cloud computing architecture to improve the DOUNG offered services. The point VI presents the cloud computing architecture of the DOUNG improved by the audio warehouse and the ATRS.

II. Basic Concepts of the Distance Education

The distance education is implemented with various techniques such as the use of the e-learning by the means of computers network. Following the evolution process, the e-learning techniques and the mobile learning [1] bring to the advent of the m-learning technology. The new concept integrates the expansion of the education including mobile learners that use a wireless network without architecture (ad hoc technology) interconnected to the initial Internet backbone. The m-learning uses a wide range of mobile devices including wireless technologies, different protocols and applications. The designed architecture operates under the service-oriented technology philosophy. The increasing number of mobile devices becomes a favorable factor for the implementation and the extension of the m-learning, particularly because of the interconnection between wired and wireless networks, and between telecommunications and computer networks.

In a second side, the technological evolution process brings to the advent of the cloud computing technique that allows to optimize the use of the computers in a network. The cloud computing provides the possibility to put in common the overall storage capacity and processing power of the computers in the network. With this technique, the network becomes a channel that carries program and data from one point to another to ensure remote collaboration between nodes seamlessly to the user. The cloud computing allows to design a configurable architecture of computing resources in the network such as storage servers, applications, and services. Following the success of the cloud computing, the technology evolution process brings to the advent of the cloud learning. The cloud learning is a service offering distance learning based on cloud computing. It uses the availability of network servers offering data that contains sources and training documents in different formats such as text, multimedia, images. The environment used to design the system is called Cloud Learning Environment (CLE). It is free of an organization and put together learners and trainers.

At the third side, the Virtual Private Network [2][3][4][5] (VPN) has emerged and becomes another technology of distance education used by many centers. The VPN technique uses several remote computers, in the image of a real private Local Area Network (LAN), but communicating by using a public infrastructure. The computers operate by creating a tunnel and imposing the constraints of (1) user authentication, (2) encryption of the data, (3) access keys management and (4) multi-protocol availability. The following protocols are used to ensure the VPN operation. The Point to Point Tunneling Protocol (PPTP) creates a virtual private network by using the Point to Point Protocol (PPP) and the Internet Protocol (IP). The Internet Protocol Security (IPsec) [6] or the Multi-Protocol Label Switching (MPLS) [7] are used at the level 3. The Secure Sockets Layer (SSL) [8] is used at level 4 and ensures the security and the confidentiality of the communication through the network. The SSL protocol uses a socket connection to associate clients to server stations. It can be used by any application for securing the traffic such as Hyper Text Transport Protocol (HTTP). Other applications are used to offer multiple choices to the learners such as File Transfer Protocol (FTP), Telnet [9], Internet Relay Chat (IRC) ...

III. The Hybrid Model of the DOUNG

A. *The DOUNG architecture*

As defined in [10], it is designed to achieve the ubiquitous faculty of the m-learning associated to the capacities offered by the cloud computing. The architecture is based on the hybrid model with set of solutions based on computers and telecommunication networks. The model interconnects teachers and learners within a course produce in a support made available for learner's access. The package 10 in 1 that summarizes the set of solutions gives different format of the link between the teacher and the learner. It includes also the format of the content of the course, the course medium and the methods of the course production. The DOUNG is designed with various method destined to make the course available (broadcast and/or storage). It is designed with the definition of many kinds of learner access (synchronous and/or asynchronous transfer files). The definition of the nature of the channel that can be used completes the description of the DOUNG services. The transmission channel consists of a backbone network connected to the Internet, accessible via Internet and GSM. The course can be subjected to immediate transmission in multimedia stream, or stored as multimedia files, treated text files, untreated text files or pdf files. The operation of the DOUNG is based on the cloud computing with the goal to achieve the balancing of the servers load in the network combines to the need to incorporate the limited capacity of mobile devices.

At the learner side, the required materials consist of a computer with Internet access or a cellular phone. Among many choices offered by the DOUNG, the learner can use multimedia applications (real time and deferred time), or file transfer mode with file displaying applications according to their format [1].

B. *The DOUNG applications*

The implementation tools of the DOUNG are previously defined in [10] with hardware, protocols and applications. The wide range of applications is given by the m-learning and is extended by the use of the cloud computing, making helpful their grouping into categories. The applications used to produce the courses are grouped in category (1). The second (2) category contains applications used to make those courses available. The applications allowing the learner to access the DOUNG services are grouped in the third (3) category. The fourth (4) category regroups applications used for the visualization of the courses by the learner and the fifth (5) category concerns the applications of exchange between teacher and learner.

The list of applications used by the DOUNG is extended because of the need to integrate the new proposed voice warehouse. As previously stated, the first list of applications of the category (1) regroups multimedia stream capture and recording applications, text processing applications, spreadsheets and pdf file generator. The category (2) still remains in multimedia recording with instant transmission server and an ftp server. Immediate multimedia viewer and ftp clients are used for the category (3). Deferred visualization multimedia client, word processing application, spreadsheets and pdf viewers are used for the category (4).

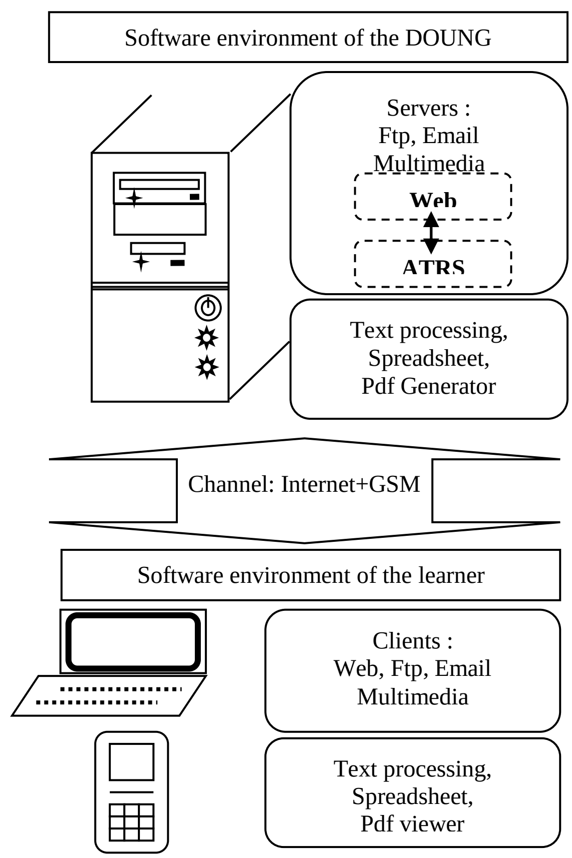


Fig. 1. The improved DOUNG software environment

The category (5) concerning applications used for the exchange between learner and teacher includes multimedia clients, multimedia servers with instantaneous transmission, Internet telephony applications, email clients and servers applications. A web server is still being suitable to be used for its multifunctional operating mode that incorporates web pages containing courses, multimedia content, ftp site, search of archives, … The following figure 1 shows the improvement of the DOUNG software environment when integrating the ATRS.

IV. The Voice Warehouse

A language can be singularly defined as a set of reserved words and a set of rules that explain how to put the language words together. The set of words define a dictionary which is usually available in a text format. Every word used to build sentences that composed a text in a particular language is provided by the dictionary. The voice warehouse is here a kind of dictionary containing the audio format of each word in a given language. A difference that can be stated is that words that have the same pronunciation generate the same audio file, even if their orthography is different.

The idea given as part of the main contribution of this paper is to improve the options offered by the DOUNG to a learner. In the improved scheme, the DOUNG provides in addition to the previous courses format, the same courses in audio format and in different language through the Advanced Text Reading System. The ATRS is implemented as a script that uses the audio warehouse containing the pronunciation of every word given by the dictionary of a selected language. The script is written to follow the words juxtaposition in a text and for each word, it will select the corresponding audio file from the warehouse. After the selection step, the audio file is played by using the loudness parameters defined later to help to take the punctuation into account. The audio flow produced is conveyed through the network (Internet + GSM) to the cell-phone of the learner as in the case of the VPN multimedia traffic between Internet and GSM. The audio flow can be also converted in a multimedia file so that a learner having a cell-phone with appropriate option (multimedia client) can download the file and play it locally. This means that another course in multimedia format can be produced by the ATRS without imposing to the teacher the constraints to record again the same course given in the text format.

Such kind of association between a word and its associated audio file can simply be made by given to the audio file, the same character composition of the word regardless to the upper or lower cases of the characters. This brings to make the script more intelligent so that it can resolve the cases of words having different orthography with the same pronunciation. A faire organization can be made in the storage of the sound in the warehouse. The audio dictionary of a language will be put into a separate directory different to the other language one. This may help to facilitate the language word selection while the language itself becomes an input parameter of the ATRS script. It can be assumed that if the text is well written in a particular language, by following the language grammar and its orthography constraints, the ATRS will also produce the correct reading of that text.

The precedent list of applications of category (1) is extended to integrate the ATRS script. The performance of the production of the sound by the ATRS is subjected to the capacity of the computer to run fast that conversion program. Another solution can be implemented used by multimedia applications. It concerns the use of the anticipation window that allows to early provide new sounds by accessing to the warehouse when the precedent audio words are being played. Thus by using the cloud computing mechanism, the task will remain on the server of the GSM provider. And then, the flow of sound provided by the ATRS will be conveyed to the cell-phone through GSM channel as in the case of the transmission of the normal audio emission. The cell-phone will receive the text, but also the sound, and the learner can choose freely to replay the sound unlimitedly, with more connection benefit in the case of a downloaded multimedia file.

Three processes are used to build the entire system. The first process is related to the building of the voice warehouse, the second process allows to implement the ATRS algorithm, and then the third process includes the running of the ATRS script by the learner.

The process of building the voice warehouse as the first step is conduct by recording an audio file for each word in a selected language. The file represents a word and is put in the warehouse. The warehouse is subdivided into folders with a separate folder for each language. A folder contains all the necessary voice files of a language that are previously recorded. Thus the whole audio warehouse becomes an audio dictionary.

V. The Advanced Text Reading System

The encoding of the ATRS script is realized by following some steps in order to facilitate its evolution through new versions. The first step allows the script to take a language as an input parameter. It sets the language folder in the warehouse as default and current folder to avoid the searching of the file path during every warehouse access. The second step of the encoding of the ATRS system brings to write the script that reads the text by following the words sequence inside the course provides in text format. For the next step, the script realizes the conversion of the extracted word into its audio format by accessing to the appropriate part of the warehouse and selecting the audio file corresponding to the word. The encoding of the ATRS script is realized within the text reading algorithm with taking care of punctuation and performing another algorithm for reading numbers. The running of the ATRS script by a learner is managed by the use of the web server.

A. *The text reading algorithm*

Let assume that the course is produced in the web page format and the ATRS system has to skim the text with the word recognition philosophy by using separators. It is useful to integrate the difference between the word of the course and the reserved words of the Hyper Text Markup Language (HTML) so that, the HTML reserved words must not be pronounced by the ATRS script. The algorithm will use the beacon format in the HTML language to distinguish the reserved words of the language and the course words. The text will be read character by character. Between HTML beacon characters starting with "<" and ending with ">", the words are ignored. Outside the range of the HTML beacon containing eventually attributes and delimited by the precedent characters, each character not includes in the list of separator is put in a string. The string is built by considering the characters between two separators. The extracted string triggered the loading and the running of the corresponding audio file from the voice warehouse. This is referenced here by the term "resolving the word sound". The text reading algorithm can be written in a low level language as the C language used in the development of the Linux version and the other open sources.

B. Integrating the punctuation in the ATRS

When reading the text, taken care of the punctuation offers the guarantee to make the reading understandable. We will consider two kinds of influence of the punctuation. The first influence is related to the break time before continuing the reading; the second one is related to the sound volume at the end of sentences. When reading the text, the ATRS will use the Normal Inter Words (NIW) interval time that allows the listener to distinguish the "space" separator between words. The speed level of reading can be set according to a value evolving inside an interval given by [LS, HS], where LS is the allowed Low Speed value and HS the allowed High Speed value. These parameters are paramount important for the regulation of the speed of the words resolution in the warehouse. They must be chosen according to the minimum time that can take the system to resolve the word in the warehouse, even if an anticipation window can be used at the convenience of the implementation. When encountering a coma (","), the ATRS will use the Short Inter Words (SIW) interval time. When encountering a full stop character ("."), the algorithm will use the Long Inter Words (LIW) interval time. When playing the pronunciation of the words extracted from a text, the algorithm will use a Normal Loudness Level (NLL). But the pronunciation of the word that comes before the full stop will use a Decrease Loudness Level (DLL) from the NLL level to the Limit of Hearing Level (LHL). The association of the NLL and the LIW at the end of a sentence can give the pronunciation of the question mark "?". Thus, LS, HS, NIW, SIW, LIW, NLL, DLL, LHL are classified in the range of the algorithm parameters that can be adjusted to bring an improved listening quality.

C. The left-to-right reading numbers algorithm

A particular algorithm must be used to read numbers differently to the reading of strings. A string is delimited by separators and its audio format is in a single file. A number is also a string delimited by separators but need to be read digit by digit and by mixing the ten powers showing the position of the digit inside the number. When reading a number, also sub numbers that are in the language base must be detected. Every language has its own base containing digits (single numbers) and sub numbers composed with digits and having their own name. The ten powers are particular sub numbers that are in the base and are usually used to determine the position of the precedent number in the sequence forming one string number. For example, 2345 can be read as follow: two thousand three hundred forty five. Thousand and hundred are the ten powers and two, three and five are digits, then forty is a sub number.

For recognizing a sub number when reading a string number, the script must take into account the current digit and the number of the remaining digits that follows it. If the system crosses a digit that begins a basic sub number, and if the number of symbols (digits) that follows the current digit allows to recognize the sub number, the system will get the other next digits before resolving the word sound. The number of symbols that follows the current digit is also used to determine the time at which the correspondent ten power must be used. For example, for "23", the system recognizes first the digit "2" and then detects that the number of the remaining digits allows to recognize a sub number. The script waits to the get the next digit before resolving the word sound. The algorithm will resolve "20" before "3". It will pronounce after getting the digit "3": twenty three. But in the case of the number "230", after recognizing the digit "2", the number of the remaining digits can't allow to recognize a sub number. Then the system will resolve "2" with using the ten power "100" and adding "30". It will pronounce: two hundred thirty.

A string that consists of characters and numbers is assumed to be a string and can be read in two ways: (1) by spelling the characters that compose the string or (2) by pronouncing the substring formed by the characters and spelling the integrated number digit by digit. For example, the string "MAN123" can be read "M-A-N-one-two-three" or "MAN-one-two-three".

The need to integrate the reading of numbers brings to the completion of the warehouse by including digit sound files, ten powers sound files, and by extension, all the sub numbers that are in the numeric base and that have their own name. For example for the English language, the base consists of the digits from zero to nine, the sub number from eleven to twenty, and thirty, forty, fifty, sixty, seventy, eighty, ninety, hundred, thousand, million, billion.

D. Reading extra language words

In a text, all the words are not necessary given by the language dictionary because of names of objects or technical names that can be used. Two kinds of solution can be implemented when during the resolution of the word in the audio warehouse the ATRS doesn't find the correspondent audio file. The first solution useful for the words in capital letter is to spell the word character by character. This means that the resolution is done for every character in the audio warehouse; not for the entire word. For example, "HTML" will be pronounced "H-T-M-L". The second solution is to put the text under test after producing the course so that the script can early detect the additional words to be put in the language dictionary and thus, in the audio warehouse. This will let the dictionary becoming richer. Anyway, it is assumed that after producing a course, the author must perform the ATRS script and listen to the reading of the text to detect anomalies. It is also possible for the author to adjust the text reading parameters to make his message understandable at his convenience. The access of the text reading parameters given as input to the ATRS script can also be let at the convenience of the learner.

E. Integrating new other language subtleties

Dependent to the robustness of the ATRS, the DOUNG can offer many courses in many languages, regardless to the evolution of the content of the course. This is made possible because of the separation between a course and its content provided in text format in one side, and in the other side the audio warehouse containing audio files of the basic words of a language. The ATRS is used to read any course provide in the language. This contribution is different to the case by which a particular course is recorded in audio format as in the case of the course provide with the multimedia captors and recorded in an audio file. In that case, when the content of the course changes another audio file must be recorded. The audio warehouse offers flexibility to teachers without the constraints to record any course or to record again when the content of the course changes. The ATRS must be designed as an evolving algorithm because every language has its own subtleties. Thus, it is paramount important to let the possibility to make the ATRS system becoming more robust among versions so that it can integrate other particular languages pronunciation of words. The following figure 2 shows the operating mechanism of the ATRS script.

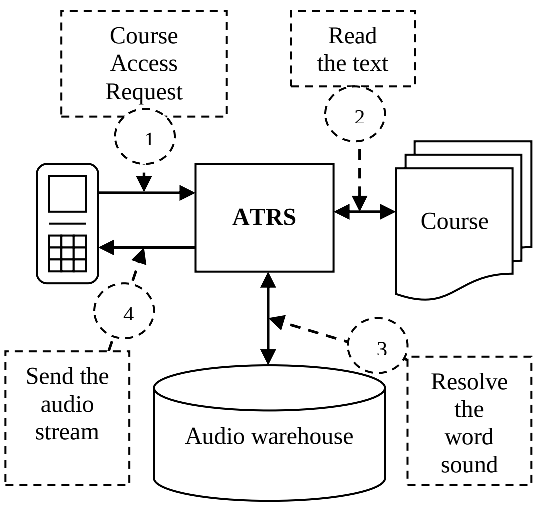


Fig. 2. Interactions between the learner, the ATRS and the courses

VI. Improving the Cloud Computing Architecture of the DOUNG

The DOUNG services are designed to operate on a cloud computing architecture in a completely transparent manner to the learner. The cloud computing is integrated to achieve a minimum level of quality of service when the learner is monitoring the courses. The Sotfware as a Service (SaaS) part of the cloud computing in the DOUNG architecture allows to subdivide the applications in elementary modules freely assembled. A web service is suggested to be used for assembling a course in image format, video format, and fixed text format or by extracting information from a database. For an advanced use of that component of the cloud computing, the ATRS script will accompany a web page by providing the audio format of the transferred text on the learner's device. The same initial advantage is maintained, that is to fit the case of GSM supports more suitable for the audio transmission. This additional ATRS script provides help for the cell devices by offering a solution to the problem inherent to their limited capacity of installing and running applications.

The Data as a Service (DaaS) of the cloud computing used by the DOUNG consists initially of available courses in a warehouse. The warehouse is implemented on different servers to resolve the limited capacity problem of data storage of the cellular devices. The contribution of the ATRS system brings to the improvement of the new version of the DaaS part of the services offered by the DOUNG. The courses warehouse in HTML format (text, image, video) is having as neighbor, the audio warehouse that is used to generate the voice format of courses in different languages. The same transparent location is maintained within a collaborative exchange between the DOUNG servers and all GSM partners.

The Platform as a Service (PaaS) part of the DOUNG services provides a support for processing. In the case of the ATRS, the PaaS is improved by the offered option to the learner to run the script of the translation of a course from its text format to its equivalent in audio format. That audio format can be easily conveyed through the GSM network. By considering the limitation of the text traffic in the GSM network, the audio format is assumed to provide an appropriate framework for the learner to follow his courses. The improved PaaS service by the integration of the remote ATRS system will help to overcome the problem of the limited processing capabilities inherent to mobile devices.

The Infrastructure as a Service (IaaS) is provided in [10] with the interfacing between the web service and other scripts. The example of CGI scripts is given which are written in low level programming languages and which execution extends to files access. The case of ATRS script is illustrative of the implementation of that part of the DOUNG service for the need of reading the text file, the need of recognizing words and selecting the appropriate audio file in the voice warehouse. The illustration is extended to the running of the audio file and its transmission through GSM channel.

VII. Conclusion

The e-learning, the m-learning, the VPN, and the cloud computing are techniques used in an integrated architecture using Internet and GSM for the advent of the DOUNG with the goal to cover large world areas. The initial model on which the DOUNG is operating let appears some complementary work of re-recording lessons according to the language of learners or according to some changes that occurs in recorded sessions. To overcome these limitations, the new model integrates the use of an audio warehouse with its ATRS exploitation script. An audio warehouse is an audio dictionary of a chosen language, recorded in a particular folder with a fair organization associating each language to a specific folder. The ATRS script is designed to read a course produced in the HTML format by extracting words, and accessing to the language audio warehouse. It then gives the correct pronunciation of the text. The system conveys the voice stream through the network. That format is more suitable for the GSM, while the cloud-computing gives facility to cell-devices used by to follow a course according to his own budget and time table. An incoming work to be conduct is to integrate many languages constraints in the ATRS. The second work is related to the adding of the language translation module. Then, the performances of the whole system must be evaluated under simulations with specific criterions.

References

[1] LeCavé, A.D. Salamin, "Mobile Learning : Les avantages du papier virtuel," FI 1, 3 Février 2004, pp.7-9.

[2] H. Jeng, J. Uttaro, L. Jalil, B. Decraene, Y. Rekhter, R. Aggarwal, "Virtual Hub-and-Spoke in BGP/MPLS VPNs", October 2013, RFC 7024

[3] Y. Cai, E. Rosen, IJ. Wijnands, "IPv6 Multicast VPN (MVPN) Support Using PIM Control Plane and Selective Provider Multicast Service Interface (S-PMSI)", RFC 6516, February 2012

[4] B. Davie, F. Le Faucheur, A. Narayanan, " Support for the Resource Reservation Protocol (RSVP) in Layer 3 VPNs ", RFC 6016, October 2010

[5] E. Rosen, Y. Cai, I. Wijnands, "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs", RFC 6037, October 2010

[6] S. Frankel, S. Krishnan, "IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap", RFC 6071, February 2011

[7] K. Lam, S. Mansfield, E. Gray, "Network Management Requirements for MPLS-based Transport Networks", RFC5951, September 2010

[8] A. Freier, P. Karlton, P. Kocher, "The Secure Sockets Layer (SSL) Protocol version 3.0", RFC 6101, August 2011

[9] T. Murphy, P. Rieth, J. Stevens, "IBM's iSeries Telnet Enhancements", RFC 4777, November 2006

[10] Issoufou Tiado, H. Saliah-Hassane, "Cloud-Computing based architecture for the advent of a New Generation of Digital Open Universities in m-learning", ICEER 2013 conference Proceedings, pp572-579, Marrakesh – Morocco, July 2013

Recognition of Similar Wooden Surfaces with a Hierarchical Neural Network Structure

Irina Topalova
Faculty of German Engineering Education and Industrial Management
Technical University of Sofia, Bulgaria
Sofia, Bulgaria

***Abstract*—The surface quality assurance check is an important task in industrial production of wooden parts. There are many automated systems applying different methods for preprocessing and recognition/classification of surface textures, but in the most cases these methods cannot produce very high recognition accuracy. This paper aims to propose a method for effective recognition of similar wooden surfaces applying simple preprocessing, recognition and classification stage. The method is based on simultaneously training two different neural networks with surface image histograms and their second derivatives. The combined outputs of these networks give an input training set for a third neural network to make the final decision. The proposed method is tested with image samples of seven similar wooden texture images and shows high recognition accuracy. The results are analyzed, discussed and further research tasks are proposed.**

Keywords—recognition; preprocessing; neural network; wooden surface

I. Introduction

Many modern automated inspection systems in wooden industry are developed for inspection of the wooden surface quality. All of them have costly hardware solutions and are compatible with the specific company production equipment. The existing software products for texture recognition are not intended for implementation in common purpose programmable logic controllers widely used in technological processes control. A specific task in wooden industry is the sorting of similar tiles having identical textures but with different shades or including some defects. It is difficult to obtain high classification accuracy in this case because of the high correlated texture parametrical descriptions.

In some cases the surface textures have to be recognized in movement, because of the production process specifics. Thus the motion blur noise is added to the similarities of the textures and some other kind of noise with Gaussian or uniform distribution. Thus de-correlation of the texture descriptions in the preprocessing stage is needed, as these descriptions usually feed the recognition structure itself. That is the reason to develop effective methods and algorithms aiming high recognition accuracy for different kinds of similar textures when evaluating them in production environment. As well optimal (considering the proportion between classification accuracy, calculation simplicity and cost) methods, software and hardware system solutions have to be sought, suitable for implementation in real time systems.

Taking into consideration the above discussion, a method for recognition of similar wooden surfaces applying simple preprocessing, recognition and classification stage is presented. The method is based on preliminary analyzing the correlation between different wood texture descriptions, followed by a simultaneously training two different neural networks with surface image histograms and their second derivatives. The combined outputs of these networks give an input training set for a third neural network to take the final decision. The proposed method is tested with image samples of seven similar wooden classes and many exemplars of their noisy images. The obtained results are presented and discussed.

II. Related Works

The most research in the wood product industry has been applied in the development of automatic visual inspection systems, based on the quality of the wood and the presence of defects. Usually these technologies use devices and technologies that are rather complex and expensive. A wood classification system based on several Multi-layered perceptron (MLP) neural network models has been developed and discussed in [1]. The MLP structure has been trained by the authors, using 20 input features such as angular second moment, contrast, correlation, inverse difference moment, entropy, etc., for five different image rotations [1]), extracted from the texture. In this case, the authors have obtained 1sec overall computation time and 95% accuracy. Other methods for defect detection in textured wood surfaces rely on the analysis and fusion of image series with variable illumination [2, 3]. These methods can be considered as filter-based, where the filters or feature detectors are learned from a set of training surfaces. It spends 0.5 to 1.6 seconds to process an image of 256x256 pixels and is tested in Mat Lab. The method represented in [4] is based on applying a length histogram that embodies the width and height of the grey level texture histogram and gives a maximum of 86.6% recognition accuracy. Considering the existing texture classification methods [1, 2, 3, 4] and technologies, we came to the conclusion that they are not effective for textures having identical structures, they need significant computational time, and in the most cases cannot obtain high recognition accuracy. The existing neural network software for texture recognition is applicable for simulating and testing the methods but is not intended for implementation in real time systems for control of different automated technological processes in industry.

III. Hierarhical Neural Network Recognition Structure

When investigating recognition and classification of a preliminary known texture classes, more suitable is to apply an adaptive recognition method and a supervised learning scheme, since this method gives the more accurate results. In this instance the best variant is to choose neural networks (NN) because of the good NN capabilities to adapt to changes in the input vector, to set precisely the boundaries between the classes therefore offering high recognition accuracy and fast computations in the recognition phase [5].

After choosing a NN for combination between a supervised learning scheme and utilization of the histogram parameters, the more important thing is the right choice of the input NN data. This choice is influential for the right and fast convergence of the NN, for the number of parameters in the input vector and the whole NN topology. The initial choice of variables is guided by intuition. Next the number of NN input parameters has to be optimized, developing a suitable method for their reduction aiming to preserve only the informative parameters without loose of any information useful for accurate class determination. The preprocessing stage in the recognition systems with visual image acquisition is intended for image quality enhancement, for feature extraction and constituting a feature parametrical vector. This vector is provided for feeding the chosen recognition structure with input data. Some important requirements to the input vector are: precisely description of the class using distinguishing features, high input vector correlation between different samples of the same class and high de-correlated input vectors representing different classes. Before taking the decision what kind of appropriate texture recognition method to apply, it is necessary to evaluate the similarity i.e. the correlation between the estimated description vectors [6]. As the histograms give an integral characteristic of the texture image, they are very appropriate to be used as initial parametrical description vectors of the image.

A. *The proposed method*

The proposed method aims high recognition accuracy of similar wooden surfaces applying simple preprocessing to obtain different input parametrical vectors for feeding two neural networks (NN) on the first recognition stage. The first stage NN outputs constitute the input vector for feeding the second stage NN, designed for making the final decision. Considering the real-time working of the system, the influence of the production technology specifics is reflected in the chosen input training set, i.e. motion blur and Gaussian noise are added to the images. Many recent developed methods use texture histograms, DCT or Wavelet transform over the texture histogram as NN input training sets [6, 7].

The proposed method here combines two different input training vectors aiming to reflect simultaneously the translations or stretches along the argument axis together with changing the function values. The translation of the histogram because of brightening or darkening of the image and the histogram stretches because of adding motion blur are the most possible reasons for histogram changes along the argument (grey levels) axis. As the image histograms are an integral presentation of the image texture and reflect the above mentioned image changes, they could be used as input training vectors for a recognition structure. The first and second derivatives of the image histogram represent the changes in the function values, i.e. in the height of the grey level texture histogram for each grey level point. The present research shows that the second derivative over the image texture histogram meets better the requirements for low inter class correlation, in comparison to other previous tested by the author NN input training sets [6, 7, 8]. The NNs are capable of setting precisely the boundaries between overlapping classes in the parametrical feature space. But feeding the NN with de-correlated input vectors already in the preprocessing stage, would give assistance to the training process and respectively to the recognition accuracy.

The proposed recognition structure is shown in Fig. 1. NN1 is trained with a set of histogram values H(g) representing different samples of the examined texture classes for each grey level g. NN2 is trained with a set of second derivative over the histogram values $d^2H(g)/dg^2$ representing different samples of the same texture classes. The outputs of NN1 and NN2 constitute the input training vector for the third NN3 which is intended to make the final decision, increasing the recognition accuracy. The proposed method and the hierarchical recognition structure was trained and tested with many samples of seven similar texture images with their species numbers, shown in Fig.2.

B. *Preprocessing stage*

In the preprocessing stage the correlation of different texture image parametrical descriptions between each two species was calculated. The inter class correlation coefficient of classes (species) i and j is r_{ij} and was calculated according to (1), where $H_i(g)$ is the histogram value for grey level g and $\overline{H}_i$ is the mean of all histogram values for class *i*. Next the coefficient r_{ij} was calculated when the parametrical descriptions of all species are DCTs over their histograms $DCT[H(g)]$ as it is given in (2).

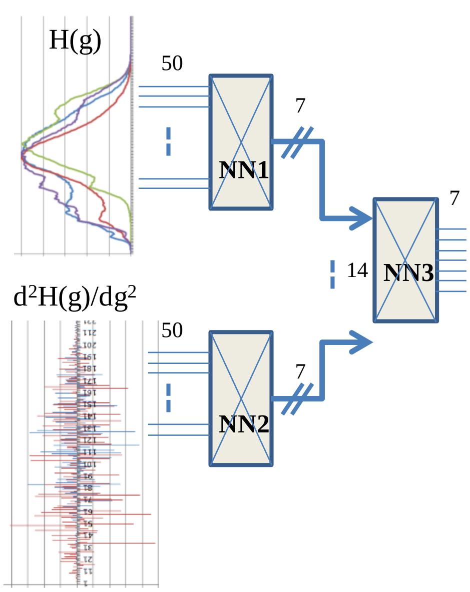


Fig. 1. Hierarchical neural network recognition structure

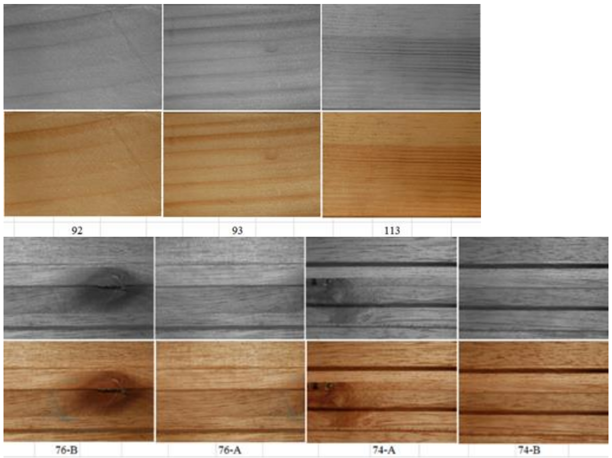


Fig. 2. Seven tested similar texture images

$$r_{ij} = \frac{\sum_{g=1}^{256}(H_i(g)-\bar{H}_i)(H_j(g)-\bar{H}_j)}{\sqrt{\sum_{g=1}^{256}(H_i(g)-\bar{H}_i)^2 \sum_{g=1}^{256}(H_j(g)-\bar{H}_j)^2}} = \frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}} \quad (1)$$

$$DCT[H(g)] = DCT(w) = c_w \sum_{k=0}^{N} H(g)\cos\left(\frac{(2k+1)}{2N}\right) \quad (2)$$

$$for\ w = \{0, \dots, N-1\},$$

$$c_w = \begin{cases} \frac{1}{\sqrt{N}}, for\ w = 0 \\ \sqrt{\frac{2}{N}}, for\ w > 0 \end{cases}$$

The correlation coefficient r_{ij} between each two texture species was also calculated for the first dH(g)/dg and second derivative $d^2H(g)/dg^2$ over the texture histogram values. Fig. 3 represents the four r_{ij} between each two wooden species over their H(g),

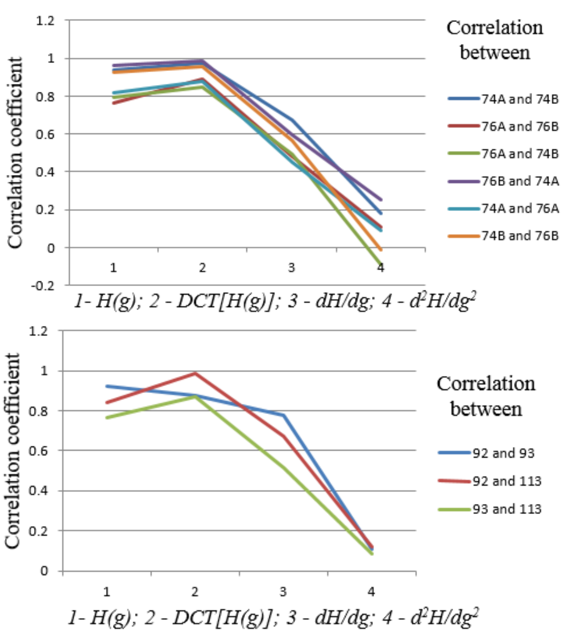


Fig. 3. Correlation coefficients between each two texture species

$DCT[H(g)]$, dH(g)/dg and $d^2H(g)/dg^2$. Obviously r_{ij} decreases considerably for the first and second histogram derivative. So it seems reasonable to use $d^2H(g)/dg^2$ as input training set because of its low inter class correlation and its capability to reflect the vertical histogram changes for each grey level g. Thus the both parametrical descriptions H(g) as an integral grey level distribution and $d^2H(g)/dg^2$ are used as input vectors for NN1 and NN2.

IV. Experiments and Results

The texture image histograms of the seven investigated species are normed through division of each histogram value by

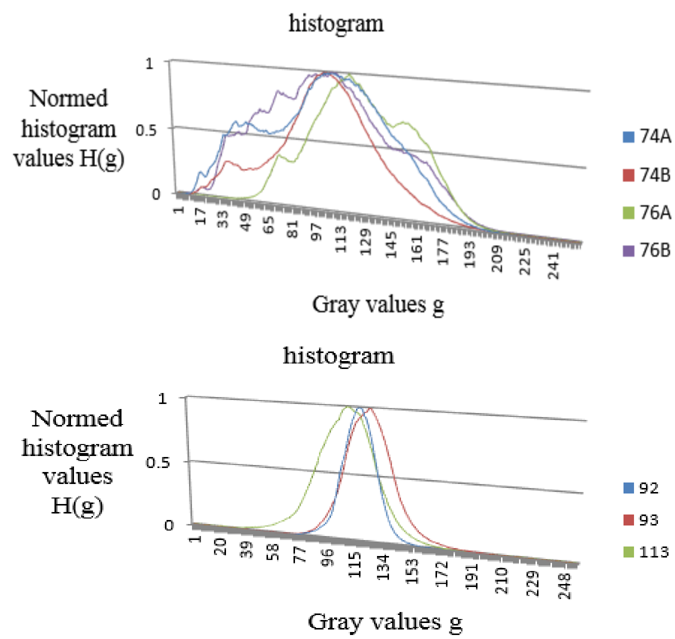


Fig. 4. Normed histogram values for the tested similar texture images

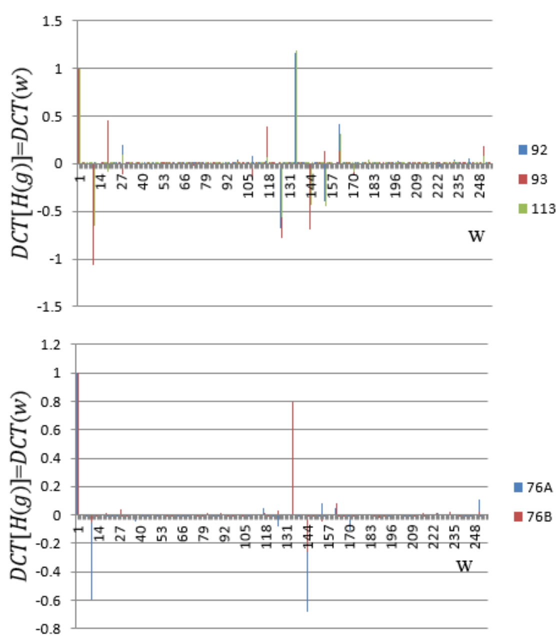


Fig. 5. Normed $DCT[H(g)]$ for the tested similar texture images

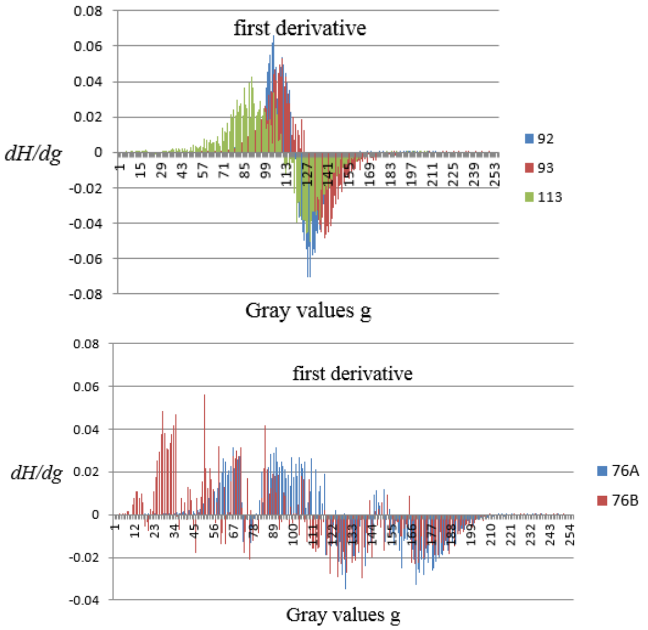


Fig. 6. Normed dH(g)/dg for the tested similar texture images

the maximum $H_{max}(g)$ value. Thus the NN1 input training set is in the argument range of the NN activation function. They are represented in Fig.4. All of the rest descriptions - $DCT[H(g)]$, dH(g)/dg and $d^2H(g)/dg^2$ are also normed in this way on account of better correlation analysis and NNs argument fitting. Fig.5 shows the normed histogram DCT[H(g)] for the tested similar texture images, respectively Fig.6 shows the normed first dH(g)/dg derivatives for the same images.

A. Training the NN structure

The images were captured with a CCD camera Nikon D7100 with CMOS 23.5x15.6 sensor and resolution of 4494x3000 pixels. The suggested hierarchical NN recognition structure is trained, tested and validated with 40 samples representing each one of the seven species given in Fig.2. Some normed d2H(g)/dg2 for the tested similar texture images are represented in Fig.7. The samples were generated adding respectively different amount - between 5 and 25 Pix - motion blur to simulate the effect of image acquisition in movement. Also 3% or 5% Gaussian noise was added to some of the images. According to the requirements of the sampling theorem [9] the number of values in H(g) and $d^2H(g)/dg^2$ were reduced to 50 points to facilitate the real-time work of the NN structure. The most frequently used proportion between training, cross validating and testing set of 60%-15%-25% of the general sample number is used in the research [10]. The 60% of the samples for each specie were randomly given to the corresponding NN1 – H(g) and to NN2 - $d^2H(g)/dg^2$. The NN1 was trained to mean square error of $\varepsilon=0.001$, NN2 to $\varepsilon=0.01$ and NN3 to $\varepsilon=0.03$. Each NN has only one hidden layer with 20 neurons. The three NNs were trained applying Backpropagation (BPG) learning algorithm since it seems to give the most promising results. *Neuro Solution software* package [10] was used, because it offers DLL export of the trained NN structure and an opportunity for implementation in different

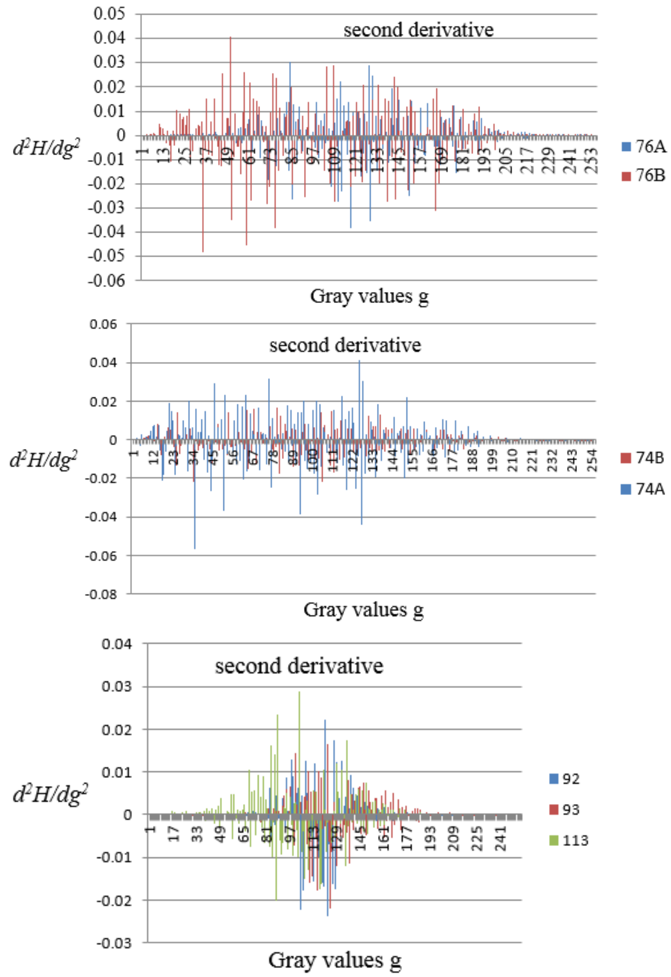


Fig. 7. Normed $d^2H(g)/dg^2$ for the tested similar texture images

Programmable Logic Controllers (PLC) for real-time work. Three different BPG algorithms [10] were probed aiming best accuracy in the recognition phase: *static* where the output of a network is strictly a function of its present input, *trajectory* where each exemplar has a temporal dimension defined by its forward period and *fixed point where* each exemplar represents a static pattern that is to be embedded as a fixed point of a recurrent network [10]. The best recognition accuracy in the recognition phase was obtained applying the static BPG learning algorithm.

B. Results and discussion

The achieved recognition accuracy when testing the trained NN structure with 10 samples (i.e. 25% of the general sample number) of each class/specie is represented in Table 1. It is visible that NN1 gives recognition accuracy between 50 and 70%, NN2 – between 70 and 90% (because the input vectors are better de-correlated), but NN3 on the last stage gives between 90% and 100%. Thus each NN contributes to the training and to the recognition accuracy in the test phase. NN3 on the second stage sets precisely the boundaries between similar texture images and classifies accurately the tested samples. Fig.8 represents NNs output values in dependence of variations of two characteristic points – $max[d^2H(g)/dg^2]$ and $VH_{(g)\ max\text{-}diff}$.

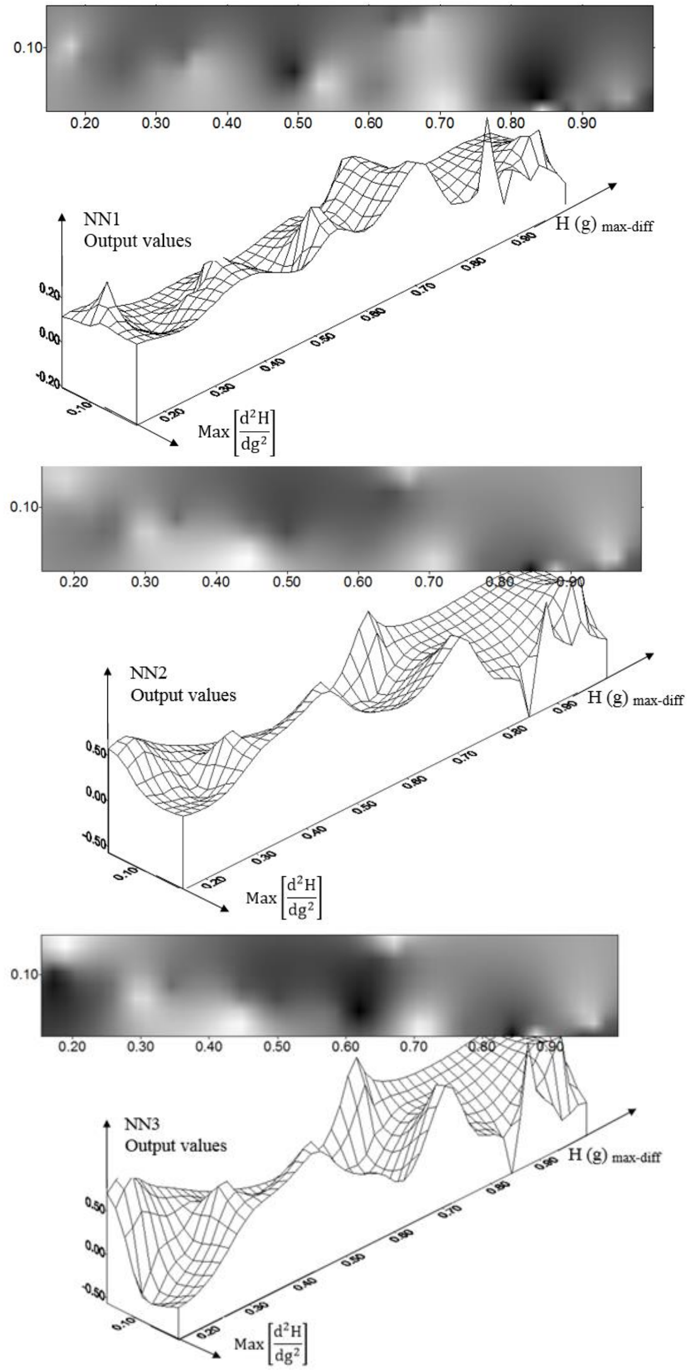
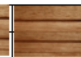


Fig. 8. NNs output values in dependence of variations of two characteristic points – Max $\left[\frac{d^2H}{dg^2}\right]$ and H (g) max-diff

These points correspond to the maximum value of the second derivative and to the histogram point where the difference between the histogram values of the tested samples is maximal. It is visible according to Figure 8, that NN3 separates and isolates better the classes in comparison to NN1 and NN2 applying closer final fitting, steeper slopes, deeper hollows and sharper boundaries.

TABLE I. RECOGNITION ACCURACY OF EACH NN IN THE STRUCTURE

Recognition accuracy [%]	Recognized Classes						
	 74-A	 74-B	 76-A	 76-B	 92	 93	 113
NN1	63	66	70	66	73	66	76
NN2	83	86	86	83	90	86	90
NN3	90	93	96	100	96	96	100

V. CONCLUSION

The main benefits of the proposed method and NN recognition structure is the achieving of high recognition accuracy by very simple preprocessing calculations. Only H(g) and $d^2H(g)/dg^2$ are calculated. Feeding different NNs with different parametrical input vectors and combining the results of the first recognition stage into input feeding vector for the second stage contributes to the precisely separating the similarities in the classes.

The obtained accuracy is between 90 and 100% even for very similar texture images. The utilized software package offers good opportunity for real-time implementation in different kinds of PLCs. To generalize the method, further research is planned with the intention of testing with much more different samples of much more texture images.

REFERENCES

[1] K. Marzuki, et al., "Design of an Intelligent Wood Species Recognition System", IJSSST, Vol. 9, No. 3, pp. 9-19, September, 2008.

[2] A.P. Grassi, et al., "Illumination and model-based detection of finishing defects", Reports on Distributed Measurement Systems, Shaker Verlag, pp. 31-35, 2008.

[3] F. Puente and K. Dostert, "Reports on Industrial Information Technology", Karlsruhe Institute of Technology, Vol 12, pp. 35-54, 2010.

[4] Hang-jun Wang, Guang-qun Zhang and Heng-nian Qi, "Wood Recognition Using Image Texture Features", PLoS One, Open Access Research,http://dx.doi.org/10.1371/journal.pone.0076101,October, 2013.

[5] M. Sartori and P. Antsaklis, A simple method to derive bounds on the size and to train multi-layer neural networks, IEEE Trans. Neural Networks, vol. 2, no. 4, pp. 467-471, 1991.

[6] I. Topalova and A. Tzokev, "Automated Texture Classification of Marble Shades with Real-Time PLC Neural Network Implementation", WCCI 2010 IEEE World Congress on Computational Intelligence, July, pp.1031-1038, ISBN 978-1-4244-8126-2/10, Barcelona, Spain, 2010.

[7] I. Topalova, "Automated Marble Plate Classification System Based on Different Neural Network Input Training Sets and PLC Implementation", IJARAI – International Journal of Advanced Research in Artificial Intelligence, May, pp. 50-56, Volume1, Issue2, ISSN: 2165-4069, 2012.

[8] I. Topalova, A. Mihailov and A. Tzokev, "Automated classification of heat resistant steel structures based on neural networks", IEEE 25-th Convention of Electrical and Electronics Engineers'08, December, pp. 437- 440, ISBN: 978-1-4244-2482-5, Eilat, ISRAEL, 2008.

[9] E. W. Weisstein, "Sampling Theorem." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/SamplingTheorem.html - last accessed on 03.09.2015.

[10] NeuroSolutions, User Manual, pp. 84-85, Copyright © 2014 by NeuroDimension, 2015.

Application of K-Means Algorithm for Efficient Customer Segmentation: A Strategy for Targeted Customer Services

Chinedu Pascal Ezenkwu, Simeon Ozuomba, Constance kalu
Electrical/Electronics & Computer Engineering Department, University of Uyo, Uyo, Akwa Ibom State, Nigeria

***Abstract*—The emergence of many business competitors has engendered severe rivalries among competing businesses in gaining new customers and retaining old ones. Due to the preceding, the need for exceptional customer services becomes pertinent, notwithstanding the size of the business. Furthermore, the ability of any business to understand each of its customers' needs will earn it greater leverage in providing targeted customer services and developing customised marketing programs for the customers. This understanding can be possible through systematic customer segmentation. Each segment comprises customers who share similar market characteristics. The ideas of Big data and machine learning have fuelled a terrific adoption of an automated approach to customer segmentation in preference to traditional market analyses that are often inefficient especially when the number of customers is too large. In this paper, the k-Means clustering algorithm is applied for this purpose. A MATLAB program of the k-Means algorithm was developed (available in the appendix) and the program is trained using a z-score normalised two-feature dataset of 100 training patterns acquired from a retail business. The features are the average amount of goods purchased by customer per month and the average number of customer visits per month. From the dataset, four customer clusters or segments were identified with 95% accuracy, and they were labeled: High-Buyers-Regular-Visitors (HBRV), High-Buyers-Irregular-Visitors (HBIV), Low-Buyers-Regular-Visitors (LBRV) and Low-Buyers-Irregular-Visitors (LBIV).**

Keywords—machine learning; data mining; big data; customer segmentation; MATLAB; k-Means algorithm; customer service; clustering; extrapolation

I. INTRODUCTION

Over the years, the increase in competition amongst businesses and the availability of large historical data repositories have prompted the widespread applications of data mining techniques in uncovering valuable and strategic information buried in organisations' databases. Data mining is the process of extracting meaningful information from a dataset and presenting it in a human understandable format for the purpose of decision support. The data mining techniques intersect areas such as statistics, artificial intelligence, machine learning and database systems. The applications of data mining include but not limited to bioinformatics, weather forecasting, fraud detection, financial analysis and customer segmentation. The thrust of this paper is to identify customer segments in a retail business using a data mining approach. Customer segmentation is the subdivision of a business customer base into groups called customer segments such that each customer segment consists of customers who share similar market characteristics. This segmentation is based on factors that can directly or indirectly influence market or business such as products preferences or expectations, locations, behaviours and so on. The importance of customer segmentation include, inter alia, the ability of a business to customise market programs that will be suitable for each of its customer segments; business decision support in terms of risky situation such as credit relationship with its customers; identification of products associated with each segments and how to manage the forces of demand and supply; unravelling some latent dependencies and associations amongst customers, amongst products, or between customers and products which the business may not be aware of; ability to predict customer defection, and which customers are most likely to defect; and raising further market research questions as well as providing directions to finding the solutions.

Clustering has proven efficient in discovering subtle but tactical patterns or relationships buried within a repository of unlabelled datasets. This form of learning is classified under unsupervised learning. Clustering algorithms include k-Means algorithm, k-Nearest Neighbour algorithm, Self-Organising Map (SOM) and so on. These algorithms, without any knowledge of the dataset beforehand, are capable of identifying clusters therein by repeated comparisons of the input patterns until the stable clusters in the training examples are achieved based on the clustering criterion or criteria. Each cluster contains data points that have very close similarities but differ considerably from data points of other clusters. Clustering has got immense applications in pattern recognition, image analysis, bioinformatics and so on. In this paper, the k-Means clustering algorithm has been applied in customer segmentation. A MATLAB program (Appendix) of the k-Means algorithm was developed, and the training was realised using z-score normalised two-feature dataset of 100 training patterns acquired from a retail business. After several iterations, four stable clusters or customer segments were identified. The two features considered in the clustering are the average amount of goods purchased by customer per month and the average number of customer visits per month. From the dataset, four customer clusters or segments were identified and labelled thus: High-Buyers-Regular-Visitors (HBRV), High-Buyers-Irregular-Visitors (HBIV), Low-Buyers-Regular-Visitors (LBRV) and Low-Buyers-Irregular-Visitors (LBIV). Furthermore, for any input pattern that was not in the training set, its cluster can be correctly extrapolated by normalising it and computing its similarities from the cluster centroids associated with each of the clusters. It will hence be assigned to any of clusters with which it has the closest similarity.

II. Literature Review

A. Customer Segmentation

Over the years, the commercial world is becoming more competitive, as such organizations have to satisfy the needs and wants of their customers, attract new customers, and hence enhance their businesses [1]. The task of identifying and satisfying the needs and wants of each customer in a business is a very complex task. This is because customers may be different in their needs, wants, demography, geography, tastes and preferences, behaviours and so on. As such, it is a wrong practice to treat all the customers equally in business. This challenge has motivated the adoption of the idea of customer segmentation or market segmentation, in which the customers are subdivided into smaller groups or segments wherein members of each segment show similar market behaviours or characteristics. According to [2], customer segmentation is a strategy of dividing the market into homogenous groups. [3] posits that "the purpose of segmentation is the concentration of marketing energy and force on subdivision (or market segment) to gain a competitive advantage within the segment. It's analogous to the military principle of concentration of force to overwhelm energy." Customer or Market segmentation includes geographic segmentation, demographic segmentation, media segmentation, price segmentation, psychographic or lifestyle segmentation, distribution segmentation and time segmentation [3].

B. Big Data

Recently, research in Big data has gained momentum. [4] defines Big data as " the word describing the large volume of both structured and unstructured data, which cannot be analyzed using traditional techniques and algorithm." According to [5], "the amount of data in our world has been exploding. Companies capture trillions of bytes of information about their customers, suppliers, and operations, and millions of networked sensors are being embedded in the physical world in devices such as mobile phones and automobiles, sensing, creating, and communicating data." Big data has demonstrated the capacity to improve predictions, save money, boost efficiency and enhance decision-making in fields as disparate as traffic control, weather forecasting, disaster prevention, finance, fraud control, business transaction, national security, education, and health care [6]. Big data is mainly characterised by three V's namely: volume, variety and velocity. There are other 2V's available - veracity and value, thus making it 5V's [4]. Volume refers to the vast amount of data in Zettabytes or Brontobytes being generated per minute; velocity refers to speed at which new data is created or the speed at which existing data moves around; variety refers to different types of data; veracity describes the degree of messiness or trustworthiness of data; and value refers to the worth of information that can be mined from data. The last V, value is what makes Big data and data mining interesting to businesses and organisations.

C. Clustering and k-Means Algorithm

According to [7], clustering is the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters). [8] opined that clustering algorithms generate clusters having similarity between data objects based on some characteristics. Clustering is extensively used in many areas such as pattern recognition, computer science, medical, machine learning. [6] states that "formally cluster structure is represented as a set of subset C=C1,……..Ck of S, such that $S=\cup_{i=1}^{k}$ Ci and $C_i \cap$Cj=∅ for $i \neq j$. Consequently, instances in S belong to exactly one and only one subset". Clustering algorithms have been classified into hierarchical and partitional clustering algorithms. Hierarchical clustering algorithms create clusters based on some hierarchies. It is based on the idea of objects being more related to nearby objects farther away [6]. It can be top-down or bottom-up hierarchical clustering. The top-down approach is referred to as divisive while the bottom-up approach is known as agglomerative. The partitional clustering algorithms create various partitions and then evaluate them by some criterion. k-Means algorithm is one of most popular partitional clustering algorithm[4]. It is a centroid-based algorithm in which each data point is placed in exactly one of the K non-overlapping clusters selected before the algorithm is run.

The k-Means algorithm works thus: given a set of d-dimensional training input vectors { $\mathbf{x}_1, \mathbf{x}_2,.., \mathbf{x}_n$ }, the *k*-Means clustering algorithm partitions the *n* training examples into *k* sets of data points or clusters $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$, where k≤n, such that the within cluster sum of squares is minimised.

That is,

$$\underset{\mathbf{S}}{argmin} \sum_{i=1}^{k} \sum\nolimits_{x \epsilon s_i} \|\boldsymbol{x} - \boldsymbol{\mu}_i\|^2 \quad (1)$$

where, $\boldsymbol{\mu}_i$ is the centroid or mean of data points in cluster s_i .

Generic k-means clustering Algorithms:

1) Decide on the number of clusters, k.

2) Initialize the k cluster centroids

3) Assign the n data points to the nearest clusters.

4) Update the centroid of each cluster using the data points therein.

5) Repeat steps 3 and 4 until the changes in positions of centroids are zero.

III. Methodology

The data used in this paper was collected from a mega retail business outfit that has many branches in Akwa Ibom state, Nigeria. The dataset consists of 2 attributes and 100 tuples, representing 100 selected customers. The two attributes include average amount of goods purchased by customer per month and average number of customer visits per month. In this paper, four steps were adopted in realising an accurate result. They include feature normalisation alongside centroids initialisation step, assignment step and updating step, which are the three major generic steps in the k-Means algorithms.

A. Feature normalisation

This is a data preparation stage. Feature normalisation helps to adjust all the data elements to a common scale in order to improve the performance of the clustering algorithm. Each data point is converted to the range of -2 to +2. Normalisation techniques include Min-max, decimal scaling and z-score. The z-score normalisation technique was used to normalise the features before running the k-Means algorithm on the dataset. Equation (2) gives the formulae for normalisation using the z-score technique.

$$x_{norm} = \frac{x - \mu_f}{\sigma_f} \qquad (2)$$

where, x_{norm} is the normalised value of x in feature vector **f**, μ_f is the meant of the feature vector **f**, and σ_f is the standard deviation of feature vector **f**.

B. Centroids Initialisation

The initial centroids or means were chosen. Figure 1 presents the initialisation of the cluster centres. Four cluster centres shown in different shapes were selected using Forgy method. In Forgy method of initialisation k (in this case k=4) data points are randomly selected as the cluster centroids.

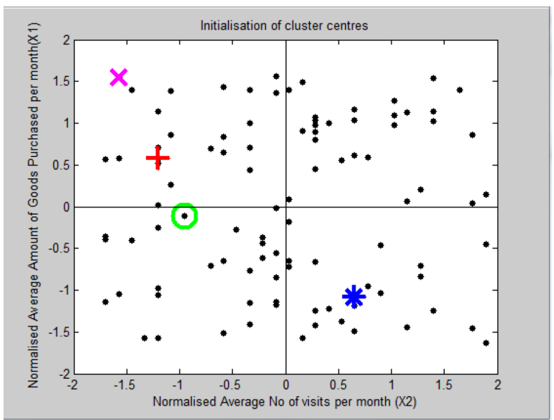


Fig. 1. The initialization stage of k-Means algorithm

C. Assignment Stage

In the assignment stage, each data point is assigned to the cluster whose centroid yields the least within cluster sum of squares compared with other clusters. That is, the square Euclidean norms of each data point from the current centroids are computed. Thereafter, the data points are assigned membership of the cluster that gives the minimum square Euclidean norm.

This has been mathematically explained in equation (3)

$$\boldsymbol{s}_i^{(t)} = \{\boldsymbol{x}_p : \left\|\boldsymbol{x}_p - \boldsymbol{\mu}_i^{(t)}\right\|^2 \leq \left\|\boldsymbol{x}_p - \boldsymbol{\mu}_j^{(t)}\right\|^2 \ \forall j, 1 \leq j \leq k\} \ (3)$$

where each data point $\boldsymbol{x}_p$ is assigned to only one cluster or set $\boldsymbol{s}^{(t)}$ at the iteration t.

D. Updating Stage

After each iteration, new centroid is computed for each cluster as the mean of all the data points present in the cluster as shown in equation (4)

$$\mu_t^{(t+1)} = \frac{1}{\left|s_i^{(t)}\right|} \sum_{x_j \in s_i^{(t)}} \boldsymbol{x}_i \qquad (4)$$

where, $\mu_t^{(t+1)}$ is the updated centroid.

Fig. 2 presents the positions of the centroids and the updated assignment of their cluster members after the 30th iteration. The each cluster members assume the same shapes as their cluster centroid. Table II shows the changes in the cluster centroids from the initialisation stage (0th iteration) to the 5th iteration.

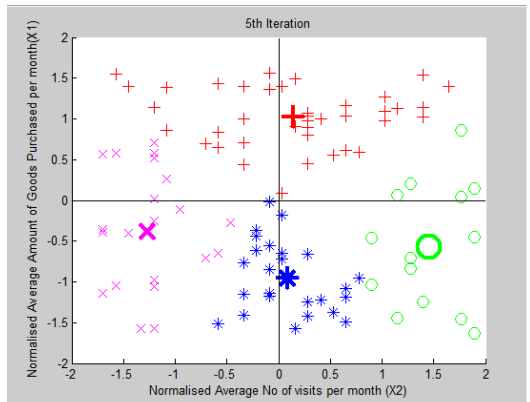


Fig. 2. Positions of the centroids and their cluster members after the 30th iteration

TABLE I. INITIALISATION AND UPDATING OF THE CLUSTER VECTORS OR CENTROIDS)

Iteration	*Cluster Centre +*		*Cluster Centre **		*Cluster Centre* **O**		*Cluster Centre* **X**	
	INTIALISED CLUSTER CENTROIDS:							
0	-0.0892	1.3654	0.6541	-1.0856	-0.2131	-0.3669	-0.2131	-0.3669
	UPDATED CLUSTER CENTROIDS:							
1	0.5656	1.0971	0.8733	-0.9508	-0.6306	-0.6728	-0.6306	-0.6728
2	0.5798	1.0456	0.9976	-0.9639	-0.5466	-0.8295	-0.5466	-0.8295
3	0.5502	1.0346	1.0376	-0.9348	-0.5600	-0.9284	-0.5600	-0.9284
4	0.5502	1.0346	1.0376	-0.9348	-0.5641	-0.9557	-0.5641	-0.9557
5	0.5502	1.0346	1.0376	-0.9348	-0.5901	-0.9894	-0.5901	-0.9894

IV. RESULTS AND DISCUSSION

The k-Means clustering algorithm converged after 100 iterations. That is, the cluster centroids became stable. Figure 3 shows the graph of the converged data points and centroids. After this, the k-Means algorithm was able to cluster almost the entire data points correctly. The centroids or the cluster vectors after convergence are:

Cluster Centre +	*Cluster Centre **	*Cluster Centre* **O**	*Cluster Centre* **X**
[-0.8325 0.9574]	[0.7403 -1.0926]	[-0.8279 -0.7217]	[0.8444 0.8412]

Each of the clusters represents a customer segment. From Figure 3, the data points at the right hand top corner represent HBRV; the data points left hand top corner represent the HBIV; the data points at the right hand lower corner represent LBRV; while those at the left hand lower corner represent the LBIV. This is clearly shown in Table II.

TABLE II. DESCRIPTION OF EACH CLUSTER IN TERMS OF THE CUSTOMER SEGMENT

HBIV *Cluster* +	HBRV *Cluster* **X**
LBIV *Cluster* **O**	LBRV *Cluster* *****

V. PERFORMANCE EVALUATION

Purity measure was used to measure the extent to which a cluster contains of class of data points. The purity of each cluster is computed with equation (5).

$$purity(D_i) = \max_j\left(P_i(C_j)\right) \quad (5)$$

Where, $P_i(C_j)$ is the proportion of class C_j data points in cluster i or D_i.

The total purity of the whole clustering i.e. considering all the clusters is given by equation (6).

$$Purity_{total}(D) = \sum_{i=1}^{k} \frac{|D_i|}{|D|} \; X \; purity(D_i) \quad (6)$$

Where, D is the total number of data points being classified.

The confusion matrix is presented in Table III.

TABLE III. CONFUSION MATRIX

Cluster	**HBIV**	**HBRV**	**LBIV**	**LBRV**	**Purity**
Cluster +	21	1	0	0	0.954
Cluster X	0	28	0	0	1.000
Cluster O	2	0	24	1	0.889
Cluster *	0	0	1	22	0.957
Total	23	29	25	23	0.950

Since, $Purity_{total}(D)$ = 0.95(from row 6, column 6 of Table 3), the clustering algorithm was 95% accurate in performing the customers segmentation.

VI. CONCLUSIONS

This paper has presented a MATLAB implementation of the k-Means clustering algorithm for customer segmentation based on data collected from a mega business retail outfit that has many branches in Akwa Ibom state, Nigeria. The algorithm has a purity measure of 0.95 indicating 95% accurate segmentation of the customers. Insight into the business's customer segmentation will avail it with the following advantages: the ability of the business to customise market programs that will be suitable for each of its customer segments; business decision support in terms of risky situations such as credit relationship with its customers; identification of products associated with each segments and how to manage the forces of demand and supply; unravelling some latent dependencies and associations amongst customers, amongst products, or between customers and products which the business may not be aware of; ability to predict customer defection and which customers are most likely to defect; and raising further market research questions as well as providing directions to finding the solutions.

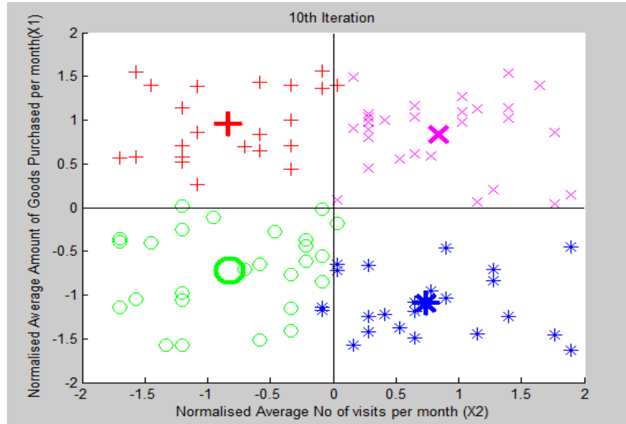


Fig. 3. The centroids converge after 100th iteration

REFERENCES

[1] Puwanenthiren Premkanth, "Market Segmentation and Its Impact on Customer Satisfaction with Especial Reference to Commercial Bank of Ceylon PLC." Global Journal of Management and Business Research Publisher: Global Journals Inc. (USA). 2012. Print ISSN: 0975-5853. Volume 12 Issue 1.

[2] Sulekha Goyat. "The basis of market segmentation: a critical review of literature". European Journal of Business and Management www.iiste.org. 2011. ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online).Vol 3, No.9, 2011

[3] *By Jerry W Thomas.* "Market Segmentation". 2007. Retrieved from www.decisionanalyst.com on 12-July, 2015.

[4] T.Nelson Gnanaraj, Dr.K.Ramesh Kumar N.Monica. "Survey on mining clusters using new k-mean algorithm from structured and unstructured data". International Journal of Advances in Computer Science and Technology. 2007. Volume 3, No.2.

[5] McKinsey Global Institute. Big data. The next frontier for innovation, competition, and productivity. 2011. Retrieved from www.mckinsey.com/mgi on 14 July, 2015.

[6] Jean Yan. "Big Data, Bigger Opportunities- Data.gov's roles: Promote, lead, contribute, and collaborate in the era of big data". 2013. Retrieved from http://www.meritalk.com/pdfs/bdx/bdx-whitepaper-090413.pdf on 14 July 2015.

[7] A.K. Jain, M.N. Murty and P.J. Flynn."Data Clustering: A Review". ACM Computing Surveys. 1999. Vol. 31, No. 3.

[8] Vaishali R. Patel1 and Rupa G. Mehta. "Impact of Outlier Removal and Normalization Approach in Modified k-Means Clustering Algorithm". IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2, September 2011 ISSN (Online): 1694-0814

[9] Md. Al-Amin Bhuiyan and Hiromitsu Hama, "Identification of Actors Drawn in Ukiyoe Pictures", Pattern Recognition, Vol. 35, No. 1, pp. 93-102, 2002.

[10] S. O. Olatunji, M. Al-Ahmadi, M. Elshafei, and Y. A. Fallatah, "Saudi arabia stock prices forecasting using artificial neural networks," pp.81–86, 2011.

[11] Q. Wen, Z. Yang, Y. Song, and P. Jia, "Automatic stock decision support system based on box theory and svm algorithm," *Expert System Application*, vol. 37, no.2, pp. 1015–1022, Mar. 2010.[Online]. Available: http://dx.doi.org/10.1016/j.eswa.2009.05.093.

[12] P.-C. Chang, C.-Y. Fan, and J.-L. Lin, "Trend discovery in financial time series data using a case based fuzzy decision tree," *ExpertSystem Application*, vol. 38, no. 5, pp. 60706080, May 2011. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2010.11.006

APPENDIX

```
clc;clf;close;clear all;
load CustData % Data file containing 100-by-2 training examples, X
%Normalisation and Selection of initial centroids
```

```
X=[(X(:,1)-mean(X(:,1)))/std(X(:,1))     (X(:,2)-
mean(X(:,2)))/std(X(:,2))];
j = 1;k=1;l=1;
i = randi(length(X));
while j==i
   j=randi(length(X));
end
while k==i|k==j
   k =randi(length(X));
end
while l==i|l==j|l==k;
   l =randi(length(X));
end
centr1 = X(i,:);centr2 = X(j,:); centr3 = X(k,:);centr4 = X(l,:);
%Initial plots of points and position of initial centroids
plot(X(:,1),X(:,2),'.k','MarkerSize',15)
hold on
plot(centr1(1),centr1(2),'+r','MarkerSize',18,'LineWidth',3)
plot(centr2(1),centr2(2),'*b','MarkerSize',18,'LineWidth',3)
plot(centr3(1),centr3(2),'Og','MarkerSize',18,'LineWidth',3)
plot(centr4(1),centr4(2),'Xm','MarkerSize',18,'LineWidth',3)
title('Initialisation of cluster centres')
xlabel('Normalised Average No of visits per month (X2)')
ylabel('Normalised Average Amount of Goods Purchased per
month(X1)')
hold off;
%Iterations to update Centroids and assign clusters members
count = 1;
while count <=10
   d1=(X-[ones(length(X),1)*centr1(1)
            ones(length(X),1)*centr1(2)]).^2;
   d2=(X-[ones(length(X),1)*centr2(1)
            ones(length(X),1)*centr2(2)]).^2;
   d3=(X-[ones(length(X),1)*centr3(1)
            ones(length(X),1)*centr3(2)]).^2;
   d4=(X-[ones(length(X),1)*centr4(1)
            ones(length(X),1)*centr4(2)]).^2;
   d11 = d1(:,1)+d1(:,2);
   d22 = d2(:,1)+d2(:,2);
   d33 = d3(:,1)+d3(:,2);
   d44 = d4(:,1)+d4(:,2);
   row1 = d11<d22 & d11<d33 & d11<d44;
   row2 = d22<d11 & d22<d33 & d22<d44;
   row3 = d33<d22 & d33<d11 & d33<d44;
   row4 = d44<d22 & d44<d11 & d44<33;
   cluster1 = X(row1,:);
   cluster2 = X(row2,:);
   cluster3 = X(row3,:);
   cluster4 = X(row4,:);
   centr1 = [mean(cluster1(:,1)) mean(cluster1(:,2))];
   centr2 = [mean(cluster2(:,1)) mean(cluster2(:,2))];
   centr3 = [mean(cluster3(:,1)) mean(cluster3(:,2))];
   centr4 = [mean(cluster4(:,1)) mean(cluster4(:,2))];
   count = count + 1;
end
% Plot the final centroids positions and cluster data points
figure; hold on;
plot(cluster1(:,1),cluster1(:,2),'+r','MarkerSize',10)
plot(cluster2(:,1),cluster2(:,2),'*b','MarkerSize',10)
plot(cluster3(:,1),cluster3(:,2),'og','MarkerSize',10)
plot(cluster4(:,1),cluster4(:,2),'Xm','MarkerSize',10)
plot(centr1(1),centr1(2),'+r','MarkerSize',18,'LineWidth',3)
plot(centr2(1),centr2(2),'*b','MarkerSize',18,'LineWidth',3)
plot(centr3(1),centr3(2),'Og','MarkerSize',18,'LineWidth',3)
plot(centr4(1),centr4(2),'Xm','MarkerSize',18,'LineWidth',3)
plot([-2 0 2],[0 0 0],'-k')
plot([0 0 0],[-2 0 2],'-k')
title('100th Iteration')
xlabel('Normalised Average No of visits per month (X2)')
ylabel('Normalised Average Amount of Goods Purchased per
month(X1)')
```