An Extended BDI Agent Architecture with Multiple Intention Reconsideration Ability in a Vessel Berthing Application

Prasanna Lokuge & Damminda Alahakoon

School of Business Systems, Monash University, Australia
Email: {prasanna.lokuge,damminda.alahakoon}@infotech.monash.edu.au

Abstract

Belief-Desire-intention (BDI) agent based systems have been implemented in many business application systems and found to have some limitations in obverting environmental changes, adaptation and learning in making rational decisions. Our paper presents a new hybrid BDI agent architecture which compares all the available intentions in the intention reconsideration process and is able to observe all the events which are related to the committed intention, before a decision is being made. Limitation in capturing of one event in the intention reconsideration process is overcome with the introduction of our extended BDI execution cycle. Further, the use of "*Knowledge Acquisition Module*" (KAM) in our proposed model improves the learning ability of the generic BDI agent. Execution of plans for a committed intention is based on the reinforcement learning techniques and Adaptive Neuro Fuzzy Inference System (ANFIS) is used in deciding the intention reconsideration of the proposed agent model. This enables the agent to interact with the environment more closely and use intelligence in making rational decisions, whose behavior may be not known at the design stage.

Introduction

Vessel berthing system is one of the complex, dynamic and large applications which require autonomy and learning to assure higher productivity and efficiency (Lokuge and Alahakoon, 2004). Our research is motivated by a vessel berthing problem faced by Container Terminal Operators in large container Ports and focuses on improving the productivity with the use of hybrid BDI agents.

Research on vessel berthing related application systems have been addressed in the literature, but to the best of our knowledge none of the papers describe the effective use of intelligent Hybrid BDI agents in assuring higher productivity and efficiency in the container terminals. Most of the papers focus on static vessel berthing system, where the main issue is to identify a good plan to assign vessels in the out harbor. (Brown et al. 1994) used integer-programming model vessel berthing, (Lim 1998) addressed the vessel planning problem with fixed berthing time, (Li et al 1998) addressed the scheduling problem with a single processor multiple jobs and assumed that vessels are already arrived, (Chia, Lau and Lim 1999) used Ants Colony optimization approach to solve berthing system minimizing the wharf length, (Kim and Moon 2003) used simulated annealing in berth scheduling. But if one could include dynamic changes during the operations in the vessel operations, use of previous knowledge and experience for the vessel planning would improves the decision making power of the system.

On the other hand, Agent systems based on practical reasoning system, which perhaps use philosophical model of human reasoning have been used in achieving optimal solutions for many business application in the recent past. BDI agent model is possibly the best known and best studied model of the practical reasoning (Georgeff, 1998) and have implementations e.g. IRMA (Bratman, 1998) and the PRS-like systems (Georgeff and Lansky, 1987) including PRS and dMARS. In some instances the criticism regarding BDI model has been that it is not well suited to certain types of behaviors. In particular, the basic BDI model appears to be inappropriate for building complex systems that must learn and adapt their behaviors and such systems are becoming increasingly important in today's context in the business applications. Another key problem in the design of BDI agent is the selection of an *intention reconsideration process* (Wooldridge, 2000), (Kinny and Georgeff, 1991). There is currently no consensus on exactly how or when an agent should reconsider its intentions (Schut and Wooldridge, 2001). A *meta-level decision theoretic approach* has been adopted (Schut and Wooldridge, 2001) to improve the agent's policy in intention reconsideration in the above paper, but still it assumes some of the information is static. Further, it is not addressed the agent behavior in uncertain or in vague environment is not addressed.

We propose an extended hybrid BDI agent model with reinforcement learning capabilities for the execution of plans in the committed intention, which essentially extend the learning and adaptability features of the current BDI agents. In this paper, we describe how dynamic changes in the environment affect the adaptive planning process in the hybrid BDI agent architecture. Proposed Adaptive Neuro Fuzzy Inference system (ANFIS) in the Hybrid BDI framework has shown improvements in learning and decision-making processes in a complex, dynamic environment. Generic BDI execution cycle has been extended to capture multiple events and compare all the possible intentions in the intention reconsideration process before a decision is being made. A trained supervised neural network in the *KAM* module is deployed in choosing the most viable intentions for an external event in the agent model (Lokuge and Alahakoon, 2004), which is not described in detail in this paper due to space limitations.

The research is carried out at the School of Business Systems, Monash University, Australia, in collaboration with the Jaya Container Terminal at the port of Colombo, Sri Lanka and Patrick Terminals in Melbourne, Australia. The rest of the paper is organized as follows: Section 2 describes the generic BDI agent model. Section 3 describes the proposed hybrid BDI architecture. Section 4 describes extended Hybrid BDI control loop. A test case is described in section 5 and conclusion is in section 6.

Generic BDI Agent Model

One of the most popular approaches to autonomous agent design is the belief-desire-intention model agency, where the notions of Beliefs, Desires and Intentions are centrally focused and often referred to as BDI agents (Rao and Georgeff, 1992).

Information about the world is described in beliefs, such as *expected time of completion of the vessel (ETC), expected time of berth of a vessel (ETB)* in a vessel berthing system. Desires indicate the set of goals that an agent could achieve at a given point in time. Agent would prefer all its desires achieved, but often desires are mutually exclusive. Therefore, agent should commit to certain desires called intentions. BDI model has pre-defined library of plans. Sequence of plans is then executed in achieving the committed intention in the agent model. Changes to the environment are reflected in terms of events. Event-queue store the sequence of events occurred during the execution of plans in the agent model. The control loop of a generic BDI agent is shown in figure 1 (Rao and Grorgeff, 1992) and (Wooldridge, 2000).

1. *$B=B_0$; $I=I_0$; π := null;*
2. *While true do*
3. *get next percept p;*
4. *B:= update beliefs ;*
5. *D:= option (B,I) ; /* get desires */*
6. *I:= select intentions (B,D,I)*
7. *π := plan (B,I) /* plan to be executed */*
8. *execute (π) ;*
9. *end while*

Figure 1: Generic BDI control loop

In line 1, the beliefs, intentions and plans are initialized. In line 2-3 agent perceives and updates its beliefs; in line 5, agent starts deliberation of possible desires and line 6, it commits to an intention to achieve. In line 7, agent generates a plan to execute. Algorithm indicated many limitations, in particular, it has assumed that the environment has not change since it observed the environment at step 3 (Wooldridge, 2000). Another limitation of the above algorithm is that the agent has over committed to its intention. i.e. all the plans which belong to the commiitted intention will be executed by the agents regardless of the envionmental changes.

Wooldridge(2000) has shown improvements to the above limitations, but does not describe how to implement the intention reconsideration process of the agent. (Schut and Wooldridge 2001) integrated the meta-reasoning in a decision theoretic model (Russell and Wefald, 1992) for deliberation process of the BDI agent architecture in the intention reconsideration process. But still there are certain limitations in the model as estimation of future environmental changes seems to be known in advance and therefore it is static.

Hybrid BDI architecture with the extended BDI control loop proposed in the paper would address the above shortcoming in complex environment. Improved characteristics of the proposed hybrid model includes, firstly, the use of supervised neural network in our *"KAM"* module deliberates and chooses most viable options that an agent could achieve. Agent then commits in achieving the most appropriate intention while monitoring other possibilities. Impacts (negative or positive) of the execution of plans are computed with the reinforcement learning techniques. After every execution, agent would look forward to see all the events happen, not only the immediate event (which is the case in the current BDI control loop) and estimates the environment change to the original state. Finally, the ANFIS is used to analyze the impact of the plan executed towards achieving the current intention and all the events occurred in the intention reconsideration process. Use of ANFIS would essentially improve the agent ability to make decisions with vague or uncertain data. Proposed hybrid BDI architecture is described in the next section.

Proposed Hybrid BDI Architecture

Interactive learning, handling environmental uncertainty and use of intelligence in making rational decisions are the primary objectives in developing hybrid BDI agents in our research. This would essentially minimize some of the limitations that exist in the current BDI agents especially in complex dynamic application systems.

Two modules proposed in the hybrid BDI architecture are shown in figure 2. *"Generic BDI Module" (GBM)* will execute the extended hybrid BDI interpreter which is described in the next section. *"Knowledge Acquisition module" (KAM)* provides the necessary intelligence for the selection of intentions in an environment, finally evaluating the appropriateness of the committed intention at the present environment. This would essentially assure dynamism in the allocation and reconsideration of committed intentions in the proposed agent model. Figure 2 shows the proposed hybrid BDI architecture.

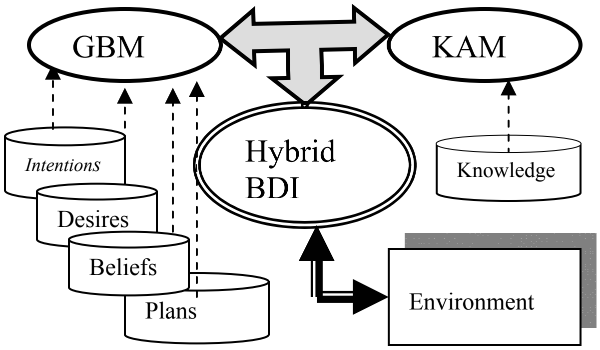

Figure 2 : hybrid BDI agent Architecture

KAM module suggested in the hybrid architecture provides the intelligence required in making decision. Neural network in the *KAM* has been trained to identify the possible desires for an external event in the environment, which is not described in this paper. ANFIS in the KAM module is used to obtain the agent's decision on intention reconsideration with the environmental changes. Agent ability to deal with partially available or vague scenarios is improved with the introduction of ANFIS. Extended hybrid BDI control loop to handle multiple intentions and multiple events is described in the next section.

Extended Hybrid BDI Control Loop

Traditional BDI agent, always observe only the next available event before it commences the intention reconsideration process. This is a limitation in the present architecture which leads to delays in making correct decisions quickly. Ability to capture all the available events related to a committed intention would essentially help agent to look forward in many steps ahead before it proceeds with the intention reconsideration process. Reinforcement learning with *n*-step Temporal Difference (TD) prediction is used to look forward and predict the effects of the future consequences towards achieving the committed intention.

Further, in the deliberation process of the present control loop, an agent will commit to an intention and ignore others, but with the environment changes, agent may have to recall previously dropped options as current one may be no longer valid at the present environment. Present BDI control loop does not support this and may cause delays in choosing a new option. Our proposed extended architecture would address the above two limitations and provide a solution to overcome the same.

Let $S=\{s_i \mid 1\le i\le n\}$ denotes *n* number of states in the environment for a committed intention. Any state s_i is described as a set of beliefs $B^I_{s,j}=\{b^I_{s,1},b^I_{s,2}\ldots b^I_{s,n}\}$ in state *s* for the intention *I*. Execution of set of plans $P^I_{s1,g}=\{p^I_{s1,s2},p^I_{s2,s3}\ldots\}$ in various states results the change of a state from one to another for a committed intention *I* as shown in the figure 3.

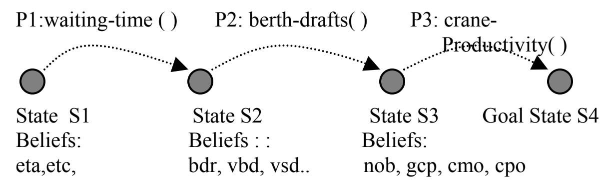

Figure 3: Plans in a committed intention *I*

Events cause the change of beliefs in different states. In the reconsideration process, it is necessary to identify the effect of the belief changes for the execution of plans. The *Belief-Impact-Matrix (BIM)* given below indicates effects of the belief changes at different states.

$$BIM=\begin{bmatrix}\alpha^{I,p}_{1,1}\ \alpha^{I,p}_{1,2}\cdots\alpha^{I,p}_{1,k}\\ \cdots\cdots\\ \alpha^{I,p}_{l,1}\ \alpha^{I,p}_{l,2}\cdots\alpha^{I,p}_{l,k}\end{bmatrix} \tag{1}$$

$\alpha^{I,p}_{i,j}$ ($0\le \alpha^{I,p}_{i,j}\le 1$) shows the impact factor or influence of the j^{th} belief in state *i* in the execution of plan *p* for the intention *I.* For example, change in *expected time of completion of a berth (etc)* does not have any effect on the execution of the plan *berth-draft( )* in states *S2,* and therefore, $\alpha^{I,p}_{i,j}$ should be zero for that instance in state *S2.* Some belief changes have higher impact on the committed intentions than others, which will be assigned values more closer to the upper bound of the $\alpha^{I,p}_{i,j}$.

Lets assume, $E^m_{(s,p)}$ and $A^m_{(s,p)}$ are the expected and actual motivation values for the execution of plans *p* in state *s,* $A^{I,t}_{(s,p)}$ is the actual distance or reward computed based on the beliefs in the environment for the plans *p* in state *s* for the intention *I* and $E^{I,t}_{(s,p)}\left(0\le E^{I,t}_{(s,p)}\le 1\right)$ is the expected distance according to the motivation value in state s for the plan *p.* Actual reward or distance due to execution of plan *p* in state *s* for a given intention *I* is given as:

$$A^{I,t}_{(s,p)} = r_{t+1}=\left\{\frac{E^{I,t}_{(s,p)}}{E^m_{(s,p)}}\times A^m_{(s,p)} \,\middle|\, s_t=s,\ p_t=p\right\} \tag{2}$$

The value of a state is the total amount of rewards or distances that an agent can expect to accumulate over the future by executing plans for an intention. Use of reinforcement learning for the execution of plans in our model is described next.

Primary objective of the use of reinforcement learning in the execution of agent plans is to learn by interaction. Results of the execution of plans are computed using the temporal difference learning which uses experience to solve the prediction problems in reinforcement learning. Execution of individual plans receives an immediate reinforcement or reward as given in the equation 2. Value function estimates the total return that can be expected to receive in the future depend on what actions the agent will take (Sutton and Barto, 1998), in this case we apply the same rule for the execution of plans for an intention.

The temporal difference learning is used to compute the expected value or distance from one state to goal state due to the execution of a plan.

$$V(s_t) \leftarrow V(s_t) + \alpha\left[A_{(s,p)}^{I,t} + \gamma V(s_{t+1}) - V(s_t)\right] \quad (3)$$

Where $\alpha\ \{0 \le \alpha \le 1\}$ is the learning rate and γ is the discount factor. Next, the agent should observe all the events occurred and compute the new state value due to change of the environment before next plan is executed. An *n-step backup* is defined to be a backup of values towards *n-step* return (Sutton and Barto, 1998). The *n-step back* method is used in capturing the value change to the state, which consider all the changes happened due to the events in the environment. Value change of state *s* due to the events is given as:

$$\Delta V_t(s_t) = \alpha\left[R_t^n - V_t(s_t)\right] \quad (4)$$

Where, R^n_t is the *n-step* return due to events observed for an intention. For example, if an agent observes that there are *n* events occurred which may have some impact on the committed intention, then equation 4 is used to compute the expected value change in state *s*. ANFIS based intention reconsideration process is then use with the above results for taking the final decision of intention reconsideration. Extended hybrid BDI control loops for multiple events/single intention and multiple events/multiple intentions are described below.

Extended BDI control loop for multiple events/Single Intention is shown in figure 4.

1. *Initialization Process; /* BIM, B, D, I .. */*
2. *While true do*
3. *α := head-of the plans (π);*
4. *execute (α);*
5. $V(s_t) \leftarrow V(s_t) + \alpha\left[A_{(s,p)}^{I,t} + \gamma V(s_{t+1}) - V(s_t)\right]$
6. *Filter-events (E^I); /* observe all related events */*
7. $\Delta V_t(s_t) = \alpha\left[R_t^n - V_t(s_t)\right]$ /*state change due to events
8. *Construct-vigilant-factor ( ); /* amount of value change */*
9. *If ((vigilant-factor > T)) then*
10. *IRF : = KAM-Intention (Ω_s^I, Φ_s^I, η_s^I, λ_s^I);*
11. *End-if;*
12. *If NOT (IRT) then*
13. *π := remaining-plans ( );*
14. *End-if*
15. *End-While*

Figure 4: Control loop for multiple event/single intention

In line 5, agent updates the state value with the execution of a plan, in line 6-7, the agent observes all the events occurred and computes the expected state value change due to environmental change. In line 9, threshold value *T* is used to control unnecessary intention reconsiderations for a small change in the value of a state. Agent *sensitivity* to environment is control with the introduction of threshold, where low values in threshold make agent more sensitive to environmental changes and vise versa. In line 10, agent uses the above linguistic vague parameters in the ANFIS to decide the intention reconsideration factor *(IRF)*. Where, Ω_s^I indicates the percentage of distance change due to the execution a plan, percentage of estimated value change due to environmental changes is given in Φ_s^I. η_s^I and λ_s^I indicate the criticality of the results obtained from executing a plan and environmental changes towards achieving the committed intentions.

We propose an extended version of BDI control loop which enable agents to handle multiple events/multiple intentions in the environment called "*global view*" i.e. while executing plans in the committed intention, agents could apply the same plan execution and environmental changes to other intentions which would have dropped at the start. This improves agent ability to rethink the use of some of the previously dropped intentions in achieving its desires. Extended control loop to handle multiple intentions is described in the following figure 5.

1. *Initialization-process ( );*
2. *While true do*
3. *calculate state value for committed intention I, $V^I(s_t)$;*
4. *calculate value due to events, $\Delta V^I{}_t(s_t)$;*
5. *For I^o=1 to n do /* for other options (intentions I^o) */*
6. *calculate state value, $V^{I^o}(s_t)$;*
7. *calculate state value due to events, $\Delta V^{I^o}{}_t(s_t)$*
8. *end (for loop)*
9. $EV^{I^o}(s_t) = MAX.\sum_{I_o=1}^{n}\left[V_t^{I^o}(s_t) + \Delta V_t^{I^o}(s_t)\right]$;
10. *Construct-vigilant-factor ( );*
11. *If ((vigilant-factor > T)) then*

12. *IRF : = KAM-Intention (* $\Omega_s^I, \Phi_s^I, \eta_s^I, \lambda_s^I, \delta_s^{I^o}$ *);*

13. *End-if;*

14. *end while*

Figure 5: Extended control loop for handling multiple intentions

In lines, 3-4 indicate the computation of state values same as given in the figure 3. In line 5, agent considers the other intentions *(I^o)* which would have ignored in the intention reconsideration process at the start. Line 6 and 7 enable agent to evaluate the opportunities to reconsider the previously dropped intentions with the current situation. Due to the environment changes observed, present intention may be no longer valid, ANFIS is used to compare the appropriateness of the committed intention with other options *(I^o_1 , I^o_2...)* which would have dropped previously. $\delta_s^{I^o}$ indicates the availability of other options in achieving the agent desire at present. This feature would enable agent to reconsider their previously dropped options if the currently committed intention is no longer valid in achieving its desires. Next section describes a test case scenario to describe the agent ability in handling multiple events and multiple intentions.

A Test Case Scenario

Lets assume, that a vessel declaration event for the vessel *Dali,"ETA-received ( )"* is received at the *JCT* terminal, Port of Colombo, which minimally includes: *vbd* ^{dali}*(vessel berth draft) = 12m, vsd*^{dali}*`(vessel sailing draft) = 12.2m, vcr* ^{dali} *(vessel crane requirement) = 13m, nob* ^{dali}*(number of boxes) = 746, eta* ^{dali} *(expected time of arrival) =0610,* etc.. *JCT* terminal has four main berths namely, *jct1*, *jct2*, *jct3* and *jct4*. Berth occupancy of the *JCT* berths are shown in the table 1.

Beliefs	*Jaami*	*Kindia*	*Dafu*	*Barrier*
Berth	Jct1	Jct2	Jct3	Jct4
brd	11.3m	12.3m	14m	14m
len	13m	13m	18m	18m
nob	337	689	612	845
Tot. hrs	12.08	31.08	29.50	54.25
gbr	43.4	27.3	35.3	29.65
etc	03:30	05:50	06:20	07:40

Table 1 : Berth occupancy details (beliefs)

When *ETA-received ( )* event is received, agent desires may be to assign the vessel to berths *jct1, jct2, jct3 or jct4*. Neural network based *"KAM"* module has chosen berths *jct2, jct3* and *jct4* as the possible options in serving the calling vessel and *jct3* as the most suitable option in assuring the highest productivity out of all the three options mentioned above. Therefore, the committed intention, *I= [Assign--berth (jct3)]*, other two possible options are I^o_1 = *[Assign--berth (jct2)]* and I^o_2 = *[Assign--berth (jct4)]*. Environmental changes are observed during the execution of plans in achieving the committed intention. In the mean time, most importantly, agent will analyze the effect of multiple events towards committed intention and compare with the other options in the intention reconsideration process. Estimated values of states for the selected desires are given in the table 2. Committed intention *I* has shown the highest value from initial state to goal state at present.

	Rewards		
Plans	I^o_1	I	I^o_2
P1 : Berth-drafts ( )	.03	.59	.59
P2 : Carne-outreach-requirements ( )	.10	1.0	1.0
P3 : Waiting-time-vessels ( )	1.0	.50	.03
P4 : Average-crane-productivity ( )	.34	.44	.50
P5 : Get-expected-operations-time ( )	.36	.47	.54
Value - state to goal D_s	1.83	3.0	2.66

Table 2 : Rewards and the estimated state value

Three different cases are taken into consideration here to investigate the agent behavior with the proposed ANFIS in the intention reconsideration process.

Case I: (local-view with one immediate event). Only the immediately available event is considered in the intention reconsideration process for the committed intention *I*.

Case II: (Local-view with many events). Effects of all the events are considered here in the intention reconsideration process for the committed intention *I*.

Case III (global-view): Effects of all the events to the currently committed intention, *I* and for all the other options *(i.e.* I^o_1 *and* I^o_2*)* are considered in the intention reconsideration process of the agent. Events occurred at a given time is given as :

E1: crane-productivity (jct3, acp=15mph) and

E2: crane-length (vcr^{dali} *=18.5m).*

One of the membership functions and the decision surface produced from the ANFIS according to the input data sets are shown in figure 6.

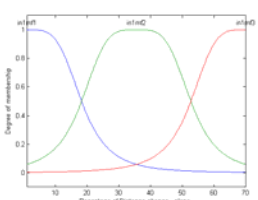

Figure 6(a) : Membership Function for percentage of Distance change due to Execution of plans.

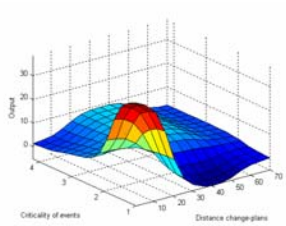

Figure 6(b): Decision Surface

Intention reconsideration decision made by the ANFIS based hybrid BDI agent in the above different cases are shown in the table 3. In case 1, environmental changes only due to *E1* are considered in the intention reconsideration process. Case II, agent observed all the events occurred which may have some effects on the committed intention *I.* Finally, in case III, agent observes all the events occurred and analyze the effects not only to the committed intention *I*, but also it looks the effect of environmental changes to the other options *(i.e.* I^o_1 *and* I^o_2) which would have dropped earlier. "NA" in the table 3 denotes "not applicable".

	Case I	*Case II*	*Case III*
Events	E1	E1,E2	E1, E2
Intentions	I	I	I , I^o_1 , I^o_2
% of distance change	8.4%	32%	32%
Available options	NA	NA	Very low
ANFIS Output – Intention Reconsideration Factor	15.1%	56.2%	23.4%

Table 3: ANFIS based output

In case I, ANFIS output indicates that agent should not drop the currently committed intention. In this case agent will only observe the effect of the environmental changes due to the immediate event occurred *("Local-view").* Case II, agent recommends to reconsider the currently committed intention as it has considered all the events occurred (environmental changes expected) in a long run. Finally, in case III,(called *Global-view*), agent look forward to capture all the events occurred and also compare the effect of the above environmental changes to the other options in addition to the committed intention. In this case, agent will observe the possibilities of adopting previously dropped options in the intention reconsideration process. If there is a better option than the currently committed one, agent will then indicate reconsideration of the current intention. In the above case III, agent has not recommended to reconsider the current intention as there are no better options available to suit the present environment.

Conclusion

In this paper we presented a new hybrid model for BDI agents that enable the use of intelligence and a mechanism in handling multiple events and intentions in the intention reconsideration process. Decision making power of the generic BDI agents is improved with the introduction of the concepts: *global-view with multiple events and intentions.* Agent ability in selecting the most appropriate intention with the environment change is improved with the introduction of the above concept.

Temporal difference learning method with *n-step* return in reinforcement learning is used to estimate the effects of the events occurred towards achieving the current intention. Events observed in the environment would have different effects at different states in the environment. Use of n-step backup method in reinforcement learning enable agent to look forward and observe all the environmental change due to events in the intention reconsideration process. Use of ANFIS in the proposed *KAM* module, then uses the knowledge in making the final decisions of whether to drop the current intention or to continue with the same. Agent ability in making decisions observing future consequences is improved with the introduction of concept of *global-view* and knowledge. Some of limitation in BDI agents, especially adaptive learning is also improved with the introduction of ANFIS.

References

Lokuge, D.P.S. and Alahakoon D. 2004. Collaborative Neuro-BDI Agents in Container Terminals, *In proceeding of the Advanced information and Networking Applications,* V1, 155-158, Japan, IEEE computer society.

Brown, G. G., Lawphongpanich, S. and Thurman, K.P. 1994. Optimizing Ship Berthing. Naval research logistics, 41, 1-15.

Lim, A. 1998. On the ship berthing problem. Operational Research Letters, V. 22, number 2-3, 105-110.

Li, C.L., Cai, X., and Lee, C.Y. 1998. Scheduling with multiple job on one processor pattern, IIE transaction 30, 45-60.

Chia, J.T., Lau, H.C and Lim, A. 1999. Ant Colony optimization for ship berthing problem. In proceeding of ASIAN99, 359-370, LNCS 1742.

Kim,, K.H and Moon, K.C. 2003. Berth Scheduling by Simulated annealing. Transportation research part B, V.37, 541-560, Elsevier Science Ltd.

Georgeff, M at al. 1998. "*The Belief-Desire-Intention Model of Agency*, Springer Publishers.

Bratman, M at al. 1998. *Plans and resource-bounded practical reasoning,* 1-22, Philosophy and (AI), The MIT Press, Cambridge.

Georgeff, M.P and Lansky A.L. 1987. Reactive Reasoning and Planning, *In proceeding of the sixth national conference on Artificial intelligence (AAAI-87),* 667-682, Seattle WA, USA.

Wooldridge, M. 2000. *Reasoning about rational agents,* The MIT Press, Cambridge, MA.

Kinny, D. and Georgeff M. 1991. Commitment and effectiveness of situated agents, *In proceeding of the 12th International joint conference (IJCAI-91),* 82-88, Sydney, Australia.

Schut, M and Wooldridge M. 2001. Principles of Intention reconsideration, *In proceeding of the AGENTS'01,* Montreal, Quebec, Canada, ACM 1-58113-326-X.

Rao, A.S and Georgeff, M.P. 1992. An abstract architecture for rational agents. *In proceeding of the knowledge representation and reasoning*, 439-449.

Russell, S and Wefald, E. 1992. Principle of meta-reasoning, *Artificial Intelligence,* 49(1-3),361-395.

Sutton, R.S and Barto, A.G. 1998. *Reinforcement Learning,- An Introduction,* The MIT Press, London.