

Değerlendirme sistemleri için melez uzman sistem yaklaşımı

Veysi ÖZTÜRK*, A. Coşkun SÖNMEZ

İTÜ Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Programı, 34469, Ayazağa, İstanbul

Özet

Sistemlerin, sentetik ortamların, insanların performansının değerlendirilmesi genellikle karmaşık olup çok zaman gerektirmektedir. Mevcut değerlendirme sistemleri belli bir alana yönelik olarak geliştirilmişlerdir ve sistemin değerlendirme sonuçlarına nasıl ulaştığını açıklamazlar. Elde edilen yeni değerlendirme bilgilerinin, değerlendirme sisteminde güncellenmesi kolay değildir. Değerlendirme süreci, uzmanlık gerektirmektedir. Fakat uzmanlar az sayıda olup, bilgilerinin bilgisayar ortamına aktarılarak daha fazla istifade edilmeleri gerekir. Bu çalışmada, değerlendirme sürecini kolaylaştıran, hızlandıran ve farklı alanlarda kullanılabilen “Genel Değerlendirme Modeli” ve “Zeki Değerlendirme Sistemi” (ZeDeS) geliştirildi. Bu kapsamda, farklı alanlardaki uzmanlardan elde edilen sezgisel bilgilerin ve farklı kaynaklardan elde edilen bilgilerin bilgisayarla değerlendirme amaçlı kullanılabilmesi için bir yöntem geliştirildi. Bu yöntemde, değerlendirme bilgileri, değerlendirme amaçlarının, değerlendirme kurallarının, ölçümlerinin, metotlarının ve parametrelerinin referans modeli olarak ifade edildi. Melez uzman sistem ve bulanık mantıktan meydana gelen “Zeki Değerlendirme Sistemi”, öğrencileri, eğitmenleri, işe başvuranları, bilgisayar tarafından meydana getirilmiş kuvvetler gibi sentetik kuvvetleri değerlendirdiği gibi gerçek sistemleri de değerlendirebilmekte olup “Genel Değerlendirme Modeli”ne ve değerlendirme ihtiyaçlarına göre geliştirildi. Değerlendirme bazı açılardan belirsizlik içerdiğinden, değerlendirmede genel çıkarım için bulanık mantıkla uzman sistemler beraber kullanıldı. ZeDeS, Hava Savunma Sistemi, öğretici performansı, pilot performansı değerlendirmesi ve eleman seçimi gibi çeşitli alanlarda ilk defa olarak kullanıldı. Makalede bir Hava Savunma Sistemi değerlendirmesinin ZeDeS kullanılarak nasıl yapıldığı ayrıntılı şekilde verilmiştir.

Anahtar Kelimeler: Zeki Değerlendirme Sistemi, yapay zekâ, melez uzman sistem, bulanık mantık.

*Yazışmaların yapılacağı yazar: Veysi ÖZTÜRK. veysi.ozturk@bte.mam.gov.tr; Tel: (262) 641 23 00 dahili: 4754.

Bu makale, birinci yazar tarafından İTÜ Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Programında tamamlanmış olan "Değerlendirme sistemleri için melez uzman sistem yaklaşımı" adlı doktora tezinden hazırlanmıştır. Makale metni 30.11.2005 tarihinde dergiye ulaşılmış, 09.02.2006 tarihinde basım kararı alınmıştır. Makale ile ilgili tartışmalar 30.06.2007 tarihine kadar dergiye gönderilmelidir.

A hybrid expert system approach for evaluation systems

Extended abstract

Evaluation of systems, synthetic environments and human performance are generally complicated and time-consuming tasks. Evaluation is needed nearly for all engineering tasks and the obstacles related with evaluation are increased proportional with complexity. Existing evaluation systems are domain dependent and do not provide explanation on how the system reaches the evaluation results. Expertise is needed for the evaluation process. Elicited new evaluation information cannot be updated to the system easily. Forming an evaluation definition is a complicated and time-consuming task. Finding out and formulating the required knowledge from the domain for which the evaluation is to be performed, is generally difficult due to lack of structured approach. It is not only important to formulate the knowledge, but finding out the right source of knowledge is also essential. Structured knowledge architecture is especially important in order to utilize evaluation knowledge automatically, especially in distributed environments. In this study, Common Evaluation Model (CEM) and INtelligent Evaluation System (INES), which simplify, speed up the evaluation process and decrease the evaluation cost, were developed. The study indicates that it is possible to put knowledge related to evaluation into a structured format. In this scope, a methodology was developed to handle the heuristic knowledge of experts from different domains and information from different sources for evaluation purposes. In this method, evaluation knowledge was represented as a reference model of evaluation objectives, production rules, measures, methods and parameters. Evaluation Objectives indicate what is going to be evaluated. Evaluation rules are criteria used to assess the collected parameters or calculated evaluation measures. Evaluation parameters are variables needed for applying rules or calculating the result of methods. The results of methods are defined as measures in order to simplify the evaluation rules and provide reusability. Evaluation methods are the algorithms for analyzing the collected parameters or / and calculating measures used in the rules. CEM shows the relation between evaluation objectives, rules, measures, methods and parameters. Using Reference Model of Evaluation Knowledge and CEM decreases the number of evaluation rules that are necessary to perform an evaluation to the related application. CEM also simplifies the representation of

evaluation knowledge. INES is a hybrid expert-fuzzy system and was developed based on CEM and evaluation needs. Before development of INES, AI techniques including expert systems, fuzzy logic, neural networks, genetic algorithms, intelligent agents and conventional programming were investigated and compared with respect to achieving high level requirements of Evaluation Systems. INES's Knowledge Base (KB) and KB Editor were developed for forming, editing and updating evaluation knowledge. INES's Inference Engine was developed for executing the evaluation definition, which includes evaluation objectives, production rules, measures, methods and parameters. Backward chaining technique was used for INES's inferencing. Some benefits of INES, which are mostly AI related, are speeding up the evaluation process, decreasing the evaluation cost, explaining the reason of evaluation results, modelling the uncertainty on an overall evaluation, providing reasoning on linguistic variables, providing a flexible structure, allow updating evaluation knowledge base without changing the source code, reducing the complexity associated with the evaluation and providing an objective and a reliable evaluation. INES was successful and was tested in the following conditions:

- *Knowledge of experts from the related domain and knowledge (or information) from the related sources for evaluation purposes are existed.*
- *Identifying evaluation criteria from the expert knowledge and information from different sources is possible.*

INES was implemented for the first time in various areas from different domains such as evaluation of Air Defence System, instructor performance, personnel selection, and pilot performance. Evaluation of an Air Defence System using INES is given in the paper. As the evaluation includes uncertainty in some aspects, Fuzzy Logic was used for reasoning. But it was realized that Fuzzy Logic could be used to perform overall performance or assessment instead of the evaluation itself for complex tasks. In other words, fuzzy logic can be more beneficial and more easily used for overall evaluation of main objective instead of all aspects of evaluation. A lot of parameters for evaluation are required and writing a lot of rules for these parameters in fuzzy logic is not an efficient way. As more rules are needed for complex systems, it becomes increasingly difficult to relate these rules to the system. Therefore, fuzzy system was used at an abstract level of evaluation.

Keywords: *Intelligent Evaluation System, artificial intelligence, hybrid expert-fuzzy system.*

Giriş

Sistemlerin, sentetik ortamların, insanların (eğitmenler, pilotlar, vb) değerlendirilmesi genellikle karmaşık olup çok zaman gerekmektedir (Öztemel ve Öztürk, 2003). Mevcut değerlendirme sistemleri, genellikle belli bir alana yönelik geliştirilmişlerdir. Sistemin değerlendirme sonuçlarına nasıl ulaştığını açıklayamazlar. Elde edilen yeni değerlendirme bilgilerinin sistemde güncellenmesi kolay değildir. Değerlendirme süreci için uzmanlık bilgisi gerekmektedir. Değerlendirme bilgilerinin bulunması ve formülize edilmesi önemli bir problem olup, değerlendiricilere gerekli değerlendirmeleri yapmalarına yardımcı olacak yapısal bir değerlendirme yaklaşımı gerekmektedir. Birçok farklı alanlar için değerlendirme metodolojisi bulunmasına rağmen hepsi için kullanılacak genel bir değerlendirme metodolojisi bulunmamaktadır (Hornung, 1995).

Bu çalışmada, değerlendirme sürecini kolaylaştıran ve hızlandıran “Genel Değerlendirme Modeli” (GeDeM) geliştirildi. Değerlendirmeyi kolaylaştıran ve hızlandıran bir modelin tanımlanması ve buna uygun zeki bir değerlendirme sisteminin geliştirilmesi zamandan ve maliyetten tasarruf sağlayacağı gibi bilgilerin yeniden kullanılabilmelerini sağlayabilir. Bu kapsamda, farklı alanlardaki uzmanlardan elde edilen sezgisel bilgilerin ve farklı kaynaklardan elde edilen bilgilerin bilgisayarla değerlendirme amaçlı kullanılabilmeleri için bir yöntem geliştirildi. Bu yöntemde; değerlendirme bilgileri, değerlendirme amaçlarının, değerlendirme kurallarının, ölçümlerinin, metotlarının ve parametrelerinin referans modeli olarak ifade edildi. Melez uzman sistem ve bulanık mantıktan meydana gelen “Zeki Değerlendirme Sistemi” (ZeDeS), “Genel Değerlendirme Modeli”ne ve değerlendirme ihtiyaçlarına göre geliştirildi. Değerlendirme bazı açılardan belirsizlik içerdiğinden, üst seviye değerlendirmedeki çıkarım (inference) için bulanık mantıkla uzman sistemler beraber kullanıldı.

Değerlendirme nedir?

Değerlendirme objeler hakkında geri besleme yapılabilmesi için sistematik bilgi edinme ve bu

bilgilere göre durumunun, kıymetinin ve uygunluğunun takdir edilmesidir. Buradaki obje bir program, teknoloji, insan, sentetik ortam ve benzerleri olabilir (Trochin, 2004). Literatürdeki değerlendirme çalışmalarından bazı örnekler aşağıda verilmiştir:

- Pilot değerlendirmesi (Shub vd., 1994),
- Simülasyon tabanlı eğitim senaryoları değerlendirmesi (Gregory, 1998),
- Uzman polis subaylarının olayları yönetim performansı (Hartley ve Varley, 2001),
- Beraber çalışan sanal ortamların performansı değerlendirmesi (Oliveira vd., 1999),
- Yüksek Seviyeli Mimari (High Level Architecture) yürütüm-süresi altyapı gerçeklemeleri değerlendirmesi (Macannuco vd., 1998),
- Yazılım değerlendirmesi (Vlahavas vd., 1998), vb.

Değerlendirmede neden yapay zekâ?

Sistemlerin, sentetik ortamların ve insanların yaptığı işler her geçen gün daha da karmaşık bir hale gelmektedir. Değerlendirme hemen her türlü mühendislik çalışmasında gerekmede ve maalesef değerlendirmenin sağlıklı bir şekilde yapılmasıyla ilgili engeller de karmaşıklıkla orantılı olarak artmaktadır. Geleneksel programlama yöntemleriyle çözülemeyen (veya çok zor çözülebilen) değerlendirmeyle ilgili aşağıdaki engeller/zorluklar bulunmaktadır:

- Değerlendirme süreci iyi tanımlanmamış olup, değerlendirme uzmanlarına göre değişmektedir. Genellikle değerlendirme, konunun uzmanı olan kişilerin subjektif gözlemlerine göre yapılmaktadır (Rigg vd., 2000).
- Değerlendirme sürecinde tecrübelerle ihtiyaç vardır. Fakat genellikle değerlendirme için az sayıda konunun uzmanları bulunmakta olup, bilgilerinin bilgisayar ortamına aktarılacak daha fazla istifade edilmeleri sağlanabilir.
- Değerlendirmeyle ilgili yeni bilgiler zamanla elde edilebileceği için değerlendirme sistemlerindeki bilgilerin güncellenmesinin kolay yapılabilmesi (meselâ, yazılımın kaynak ko-

dunun değiştirilmeden) gerekmektedir.

- Mevcut sistemler genellikle değerlendirme sonucuna nasıl ulaşıldığını açıklayamamaktadır. Değerlendirme sonuçlarına nasıl ulaşıldığının açıklanması geri besleme için çok faydalıdır. Örnek olarak, öğrencilere sadece değerlendirme sonuçlarını sağlamak yerine problemlerin (başarısızlığın) kaynağının (nedeninin) anlaşılmasını sağlamak önemlidir. Bu, gelecekte yapılacak eğitimler için öğretmenlerin ve öğrencilerin nereye odaklanacaklarına yardımcı olacaktır (Bass, 1998).

Sentetik ortamlar ve sistemler her geçen gün daha da karmaşık bir hale gelmektedirler. Manuel olarak veya geleneksel programlama ile değerlendirilmeleri zorlaşmaktadır. Mesela, öğrenciler durumlarının kolayca gözlenmesinden daha da karmaşık görevleri yerine getirmektedirler. Eğer öğretmenin gözlenmesi gereken birden fazla öğrencisi varsa aynı zamanda bütün öğrencilerinin ne yaptığını takip etmesi zordur. Buna bir çözüm, öğrencilerin yorumlama ve değerlendirmelerinin kural tabanlı bilgisayar programlarıyla yapılarak öğretmenin iş yükünün azaltılmasıdır (Drewes ve Gonzalez, 1994).

Yukarıdaki engel ve zorlukların aşılması ve kolaylaştırılması için yeni tekniklere ihtiyaç bulunmaktadır.

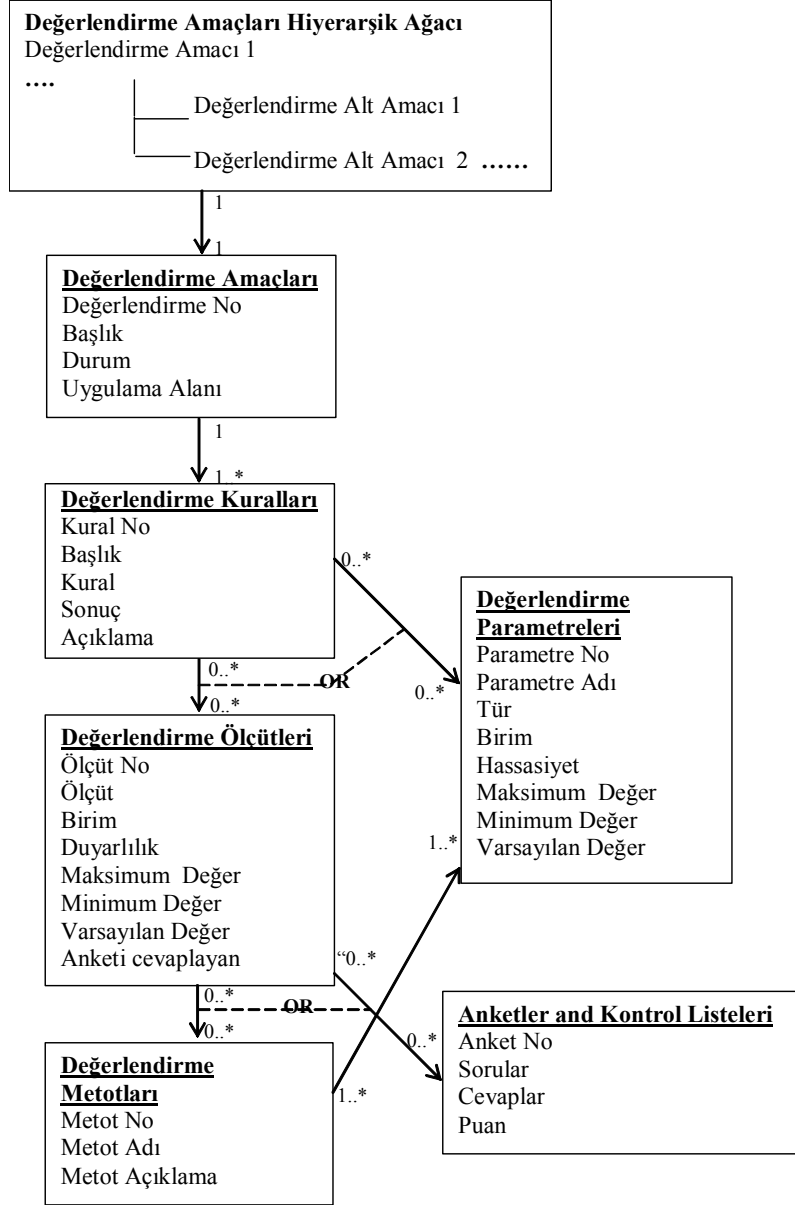
Bu çalışmada öncelikle değerlendirmeye ilgili bilgilerin modellenbildiği GeDeM geliştirilmiş, daha sonra da GeDeM'e göre değerlendirme sistemlerinin genel isteleri tespit edilmiştir. Bu istelerin yapılabilirliği, yaygın olarak kullanılan yapay zekâyla teknikleri dikkate alınarak araştırılmıştır. Değerlendirme sistemleri istelerinin, yaygın kullanılan yapay zekâ tekniklere göre yapılabilirliği Tablo 1'de verilmiştir. Bu tabloda, özellikle yapay zeka teknolojileri kullanılarak sağlanması kolay olabilecek isteler verilmiştir. Tablo 1'den görülebildiği gibi, zeki değerlendirme sisteminin geliştirilmesi için uzman sistemler ve zeki etmenler kullanılabilir, veya tabloda verilen teknolojilerin melez halleri kullanılabilir. Fakat etmene yönelik (agent-oriented) analiz ve tasarım metodolojilerinde hâlâ eksiklikler ve açık olmayan kısımlar vardır

Tablo 1. Değerlendirme istelerinin yaygın kullanılan yapay zekâ tekniklerinin yapılabilirliği

Değerlendirme Sistemlerinin İsterleri	US	BM	YSA	GA	ZE	GP
Değerlendirme süreçlerine ve kurallarına göre esnek olmak (kaynak kodların farklı uygulamalar için değişmemesi). Değerlendirme ile ilgili bilgilerin kolay değiştirilebilir bir formatta olması (Rollesbroich vd., 2003a).	√	√, fakat zor	X	X	√	X
Bilgiyi işleme zamanı (değerlendirmeyi kabul edilebilir bir zamanda yapmak)	Kısa	Kısa	Eğitim Uzun	Uzun	Kısa	Kısa
Kullanıcıya sentetik ortam, gerçek sistemler, öğretmen değerlendirmesi gibi farklı bilgi alanlarında değerlendirme yapmasına izin verme	√	√	√, fakat eğitim gerekli	Kıs men	√	√, fakat çok zor
Kayıt edilecek verilerin belirlenmesinde kullanıcıyı desteklemek	√	Kıs men	X	X	√	√
Kullanıcının ihtiyaçlarına uygun değerlendirme amacının bulunması için bilgi tabanı veya veritabanının taranması	√	X	X	X	√	√
Sisteme herhangi bir kaynak kodu değiştirmeden yeni bilgilerin eklenmesi ve mevcutların değiştirilmesi	√	Kıs men	Kıs men, eğitimle	X	√	No
Yüksek seviyeli değerlendirme ihtiyaç ve amaçlarından değerlendirme kriterleri, ölçümler, metotlar ve parametrelerin türetilmesi	√	X	X	X	√	√
Değerlendirme amaçlarındaki tutarsızlıkların engellenmesi	√	Kıs men	X	X	√	√
Sistemin sonuçlara nasıl ulaşıldığının açıklanması	√	X	X	X	√	X

US: Uzman Sistemler, **BM:** Bulanık Mantık, **YSA:** Yapay Sinir Ağları **GA:** Genetik Algoritmalar, **ZE:** Zeki Etmenler **GP:** Geleneksel Programlama

(Tidhar vd., 1999). Etmen tabanlı sistemlerin geliştirilmesi, uzman sistemlere göre daha uzun zaman almaktadır. Bunlardan dolayı bu çalışmada, değerlendirme için öncelikle yapay zekâ tekniklerinden uzman sistemler kullanılmıştır ve daha sonra uzman sistemlerin sonuçlarını en üst seviyede, genel olarak değerlendirebilen bulanık mantık modülü sisteme eklenmiştir.



Şekil 1. Genel değerlendirme modeli

Genel değerlendirme modeli

Değerlendirme bilgilerinin referans modeli olan Genel Değerlendirme Modeli (GeDeM), Şekil 1’de gösterilmiştir. GeDeM’deki değerlendirme amaçlarındaki hiyerarşik yapı, yüksek seviyeli ve alt seviyeli amaçlar arasındaki ilişkiyi gösterir. Değerlendirme amaçları, kullanıcı ihtiyaçlarından ve kullanıcı/sistem isterlerinden elde edilebilir. GeDeM’de her bir değerlendirme amacının ilgili kuralları vardır ve aynı değerlendirme kuralının farklı değerlendirme amaçları için kullanılabilmesi tekrarları önlemektedir. Benzer mantıkla, her bir değerlendirme kuralının ilgili

ölçümleri (ya da parametreleri) vardır ve aynı değerlendirme ölçümleri (ya da parametreleri) için farklı değerlendirme kurallarının kullanılabilmesi tekrarları önlemektedir. Aynı şekilde, her bir değerlendirme ölçümünün hesaplanması için ilgili değerlendirme metodları vardır ve aynı değerlendirme metodları için farklı değerlendirme ölçümlerinin kullanılabilmesi tekrarları önlemektedir. Benzer şekilde, her bir değerlendirme metodunda (ya da kuralında) kullanılan ilgili parametreler vardır ve aynı değerlendirme parametreleri için farklı metodu (ya da kuralıyla) kullanılabilmesi tekrarları önlemektedir.

Zeki değerlendirme sistemi

Zeki Değerlendirme Sistemi (ZeDeS), değerlendirme sistemleri için özel tasarlanmış bir uzman sistem ve Matlab bulanık mantık yazılımını içeren, kural tabanlı bir yazılım olup aşağıdaki fonksiyonlara sahiptir:

- Kullanıcıya yapacağı işle ilgili değerlendirme tanımlama aşamasında kullanılacak kriterler, ölçümler, metotlar, parametreler hakkında bilgi sağlar.
- Konunun uzmanı kişilerden toparlanmış bilgilerin düzenli bir formatta saklanması ve tekrar kullanılmasını sağlar. Bu, değerlendirme sürecinde kullanılacak bilgilere güveni artırır ve değerlendiriciler arasında bilgi alışverişini artırır.
- Yapılacak işle ilgili bilgi tabanında bulunan değerlendirme bilgilerinin işletilmesi (execute) ve değerlendirme sonuçlarının üretilmesi sağlar.
- Çıkarım sonuçlarının metin ve grafik olarak kullanıcıya sunar ve kaydeder.
- Çıkarıma nasıl ulaşıldığını açıklar.
- Değerlendirmedeki karmaşıklığı azaltır.
- Değerlendirme görevleri için gerekli zamanı ve maliyeti azaltır.

ZeDeS, öğrencilerin ve öğretmenlerin, işe başvuranların değerlendirilmesi, simülatör gibi sentetik ortamların değerlendirilmesi ve gerçek sistemlerin değerlendirilmesinde eğitimler gibi değerlendiriciler tarafından kullanılabilir. Hava Savunma Sistemi, öğretici performansı, pilot performansı değerlendirmesi ve eleman seçimi gibi çeşitli alanlara ZeDeS ilk defa uygulandı.

ZeDeS, bir kısmı Tablo 1’de verilen değerlendirme yazılımları istekleri, kullanıcı arayüzüyle ilgili bazı istekler dikkate alınarak geliştirildi. Borland Delphi 5 ve Matlab 6.1 kullanılarak yinelemeli (iterative) bir süreç uygulandı. Bu sürece göre, önce bir prototip gerçekleştirildi ve sürekli eklemelerle gelişimine devam edildi. Şekil 2’de gösterildiği gibi ZeDeS, üç ana bileşenden meydana gelir:

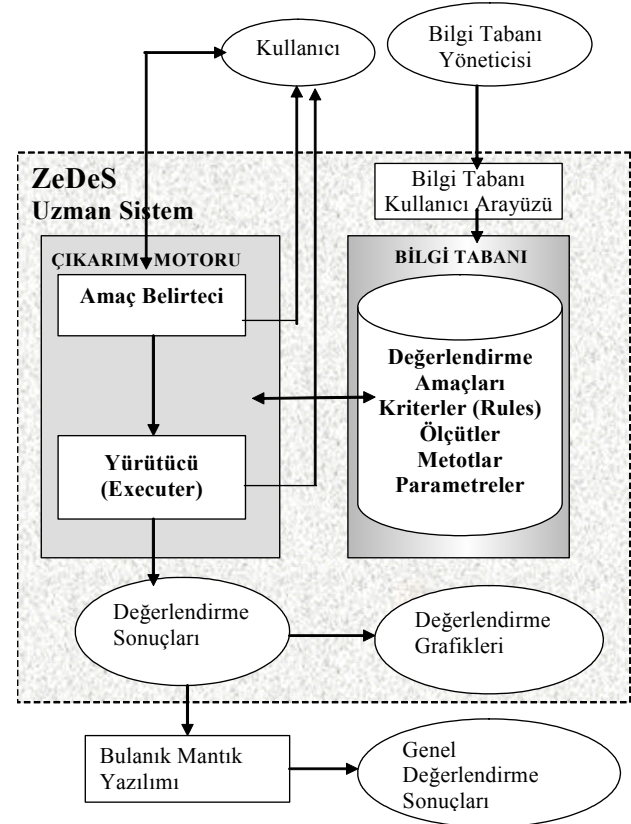
- **Uzman sistem bilgi tabanı:** Değerlendirmeyle ilgili kuralların ve referans modelin, bilgisayarın çıkarım yapabile-

ceği formata getirilmesini ve kaydedilmesini sağlar.

- **Uzman sistem çıkarım motoru:** Kullanıcı tarafından belirlenen değerlendirme amaçları için çıkarım yapmaya yarar.
- **Bulanık mantık yazılımı:** Uzman Sistem Çıkarım Motorunun sonuçlarından üst seviye değerlendirme yapılmasını sağlar.

ZeDeS uzman sistem bilgi tabanı

Uzman sistemlerin geliştirilebilmesi için konuyla ilgili alanda detaylı bilgiler gereklidir. Bundan dolayı gerekli bilgilerin toplanması, bilgisayarın anlayabileceği formata dönüştürülmesi ve yapısal bir şekilde bilgi tabanında saklanması gereklidir. Bunun için ZeDeS Bilgi Tabanı Editörü geliştirildi. ZeDeS’te bilgi gösterimi, kurallar ve hiyerarşik listeler şeklindedir. ZeDeS bilgi tabanı, uzmanların değerlendirme yapabilmesi için gerekli bilgileri ve tecrübelerini içerir. ZeDeS bilgi tabanında aşağıdaki bilgiler bulunmaktadır:



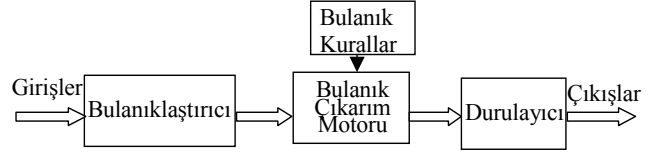
Şekil 2. ZeDeS'in ana bileşenleri

- **Değerlendirme amaç(lar)ı:** Değerlendirme amaç(lar)ının başlığı, durumu (zorunlu olup olmadığı), kullanma şekli ve varsa değerlendirme amaçlarının birbirleriyle ilişkilerini içerir.
- **Değerlendirme kuralları:** Değerlendirme kriterleriyle (hangi durumda başarılı, başarısız gibi) ilgili bilgileri içerir.
- **Değerlendirme ölçümleri:** Değerlendirme kurallarındaki hesaplanacak değerler hakkında detaylı bilgileri içerir.
- **Değerlendirme metotları:** Kurallardaki ölçümleri hesaplamak için gerekli algoritmaları içerir.
- **Değerlendirme parametreleri:** ölçümlerde veya metotlarda kullanılan değişkenler hakkında bilgiler içerir.

ZeDeS uzman sistem çıkarım motoru

Uzman sistemlerin kalbi olan çıkarım motorları, aynı zamanda kural işleyicisi olarak ta bilinirler. Bu yazılım programı, bilgi tabanındaki bilgilerden zeki kararlar çıkarılmasını sağlar. Çıkarım motorlarının bilgi tabanında ilgili kuralları bulması ve çalıştırılması için geriye zincirleme (backward chaining) ve ileriye zincirleme (forward chaining) gibi yöntemler geliştirilmiştir. Bu ve diğer yöntemler hakkındaki detaylı bilgiler için Russell ve Norvig (1995)'a bakınız. ZeDeS'te de kullanılan geriye zincirleme yönteminde, çıkarım motoru bir değerlendirme amacından başlayarak amacı karşılayan bilgileri ve gerçekleri arar. Çıkarımın yönü amaçtan ilgili gerçeklere doğrudur.

ZeDeS çıkarım motorunda, kullanıcı öncelikle konuyla ilgili anahtar kelimeler girerek, bütün bilgi tabanının derinlik- öncelikli (depth-first) arama algoritmasıyla ve veritabanı sorgulamalarıyla aranmasını sağlar. Arama sonuçları, kullanıcıya hiyerarşik şekilde ve ayrıntılı açıklamalarla sunulur. Kullanıcı, ilgi duyduğu amaçları sonuç listesinden seçtikten sonra, çıkarım motoru konuyla ilgili kuralları, metotları, ölçümleri, parametreleri ve bunları birbirleriyle ilişkilerini bulur. Bundan sonra çıkarım motoru, uygulamadan alınan verilerle ilgili kuralları, metotları, ölçümleri, parametreleri kullanarak değerlendirme



Şekil 3. Genel bir bulanık sisteminin bileşenleri

sonuçlarını üretir ve istenirse sonuçları grafik olarak gösterir.

ZeDeS bulanık mantık yazılımı

Bulanık mantık, belirsizlik durumlarında ve karmaşık problemlerin çözümünde kullanılabilen, üyelik fonksiyonlarına ve bulanık kurallara göre çalışan bir yöntemdir. Bulanık mantık, uzman sistemlerin de otomatik çıkarımda kendisini kanıtlamış yeni bir disiplindir (Konar, 2000). Genel bir bulanık sisteminin ana bileşenleri Şekil 3'te gösterilmiş olup şunlardır:

Bulanıklaştırıcı: Rakamsal verileri bulanık kümelere adresler.

Bulanık Kurallar: Girişleri çıkışlara birleştiren Eğer (If).....öyleyse (then) formatındaki kurallardır.

Bulanık Çıkarım Motoru: Bulanık girişlere göre bulanık çıkışları bulanık kurallarla üreten mekanizmadır.

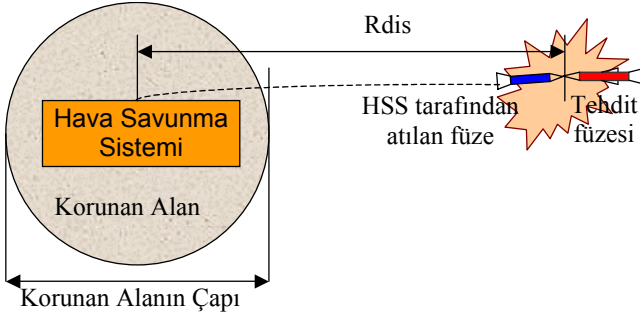
Durulayıcı: Bulanık kümeleri kesin rakamsal verileri adresler.

Şekil 3, aynı zamanda ZeDeS bulanık mantık yazılımının ana bileşenlerini gösterir. ZeDeS bulanık mantık, ZeDeS uzman sistemin sonuçlarının genel değerlendirilmesinde kullanılmıştır. ZeDeS bulanık mantığın giriş değerleri, ZeDeS uzman sisteminin sonuç değerleridir.

Hava savunma sistemi değerlendirmesi

Bu bölümde, bir Hava Savunma Sisteminin (HSS) değerlendirmesinin ZeDeS kullanılarak nasıl yapıldığı anlatılmıştır. HSS, bir bölgeyi bütün hava tehditlerinden ve özellikle füzelerden koruyan Şekil 4'te gösterilen sistemdir. HSS'nin performansı, tehditleri durdurma kabiliyetiyle belirlenir. HSS'nin bir füzeyi durdurma kabiliyetinin değerlendirilmesi için aşağıdaki alt amaçlar belirlenmiştir:

1. HSS'nin tehdit füzelerini vurma oranının değerlendirilmesi



Şekil 4. Hava savunma sistemi

2. HSS'nin reaksiyon zamanının değerlendirilmesi
3. HSS tarafından korunan alandaki hasarın değerlendirilmesi
4. Genel değerlendirme

HSS'nin performansı, çeşitli ölçümleri kullanan kuralların uygulanmasıyla değerlendirilir. HSS için kullanılan ölçümler vurma_oranı, reaksiyon_zamanı, korunan_alandaki_hasar_oranı ve genel_performanstır.

Vurma Oranın değerlendirilmesi için aşağıdaki kurallar kriter olarak kullanılmıştır:

- Eğer vurma_oranı %85'ten büyükse veya eşitse, o halde vurma oranı yeterlidir.
- Eğer vurma_oranı %85'ten küçükse, o halde vurma oranı yetersizdir.

Buradaki vurma oranının değerlendirilmesi için kurallarda %100 vurma oranı zorunlu değildir, çünkü HSS'ler gerekirse hedefi vurmak için ikinci bir füzeyi gönderirler. Vurma_Oranı için kullanılan metot ise;

$$\text{Vurma_Oranı} = (\text{vuran füzelerin toplamı} / \text{fırlatılan füzelerin toplamı}) * 100$$

HSS'lerin önemli diğer bir karakteristiği ise reaksiyon zamanıdır. Reaksiyon zamanı, tehdit nesnesinin HSS radarında tespitinden karşı füze fırlatma zamanına kadar olan süredir.

Bir füzenin korunan alanda patlama uzaklığına göre meydana getirdiği hasar oranı da HSS'nin değerlendirmesinde kullanıldı. Korunan alanın ve füzenin iki boyutlu modeli Şekil 5'te gösterilmiştir.

Burada, R korunan alanın merkezi ile patlayan füzenin merkezi arasındaki uzaklığı, R_{SA} korunan alanın yarıçapı, R_{CM} patlayan füzenin hasar meydana getirdiği alanın yarıçapı, α ve β ise korunan alanın ve füzenin daire parçalarını hesaplamak için kullanılan açılardır.

Korunan alandaki hasar oranını değerlendirmek için şu kurallar kullanılmıştır:

- Eğer R, $(R_{SA}+R_{CM})$ 'dan büyükse, o halde korunan alan hasarsızdır.
- Eğer R, R_{SA} 'dan küçükse, o halde korunan alan hasarlıdır.
- Eğer Hasar_Oranı %10'dan küçükse, o halde korunan alandaki hasar kabul edilebilir seviyededir.
- Eğer Hasar_Oranı %10'dan büyük veya eşitse, o halde korunan alandaki hasar kabul edilebilir seviyededir.

Hasar_Oranı hesaplanması için,

Hasarlı Alan=

$$2 * (\alpha \text{ açılı çember parça alanı} - \text{üçgen alanı}) + 2 * (\beta \text{ açılı çember parça alanı} - \text{üçgen alanı}) \quad (1)$$

$$= 2 * ((\pi R_{SA}^2 \alpha / 2\pi) - (R_{SA}^2 \sin \alpha / 4)) + 2 * ((\pi R_{CM}^2 \beta / 2\pi) - (R_{CM}^2 \sin \beta / 4)) \quad (2)$$

$$= \pi R_{SA}^2 \alpha - (R_{SA}^2 \sin \alpha / 2) + R_{CM}^2 \beta - (R_{CM}^2 \sin 2\beta / 2) \quad (3)$$

$$\text{Hasarlı Alan Oranı (\%)} = 100 * (\text{Hasarlı Alan} / \text{Korunan Alan}) \quad (4)$$

$$= 100 * (\pi R_{SA}^2 \alpha - (R_{SA}^2 \sin \alpha / 2) + R_{CM}^2 \beta - (R_{CM}^2 \sin 2\beta / 2)) / \pi R_{SA}^2 \quad (5)$$

$$= 100 * (\alpha - (\sin 2\alpha / 2) + (R_{CM} / R_{SA})^2 (\beta - \sin 2\alpha / 2)) / \pi \quad (6)$$

Burada

$$R_{CM}^2 = R_{SA}^2 + R^2 - 2R_{SA}R \cos(\alpha) \quad (7)$$

Kosinüs Kuralı

$$\alpha = \arccos((R_{SA}^2 + R^2 - R_{CM}^2) / 2R_{SA}R) \quad (8)$$

$$R_{SA}^2 = R_{CM}^2 + R^2 - 2R_{CM}R \cos(\beta) \quad (9)$$

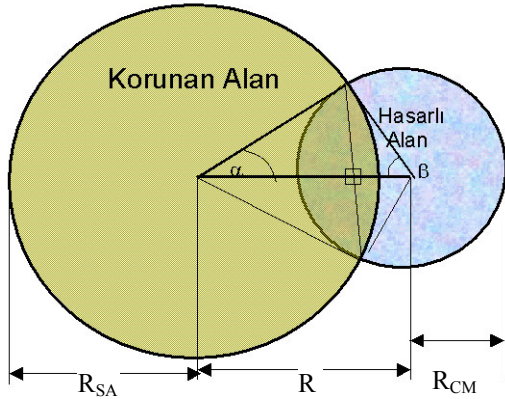
$$\beta = \arccos((R_{CM}^2 + R^2 - R_{SA}^2) / 2R_{CM}R) \quad (10)$$

Yukarıdaki kural ve metotların uygulanmasıyla elde edilen ZeDeS'in HSS değerlendirme sonuç ekran görüntüsü Şekil 6'da verilmiştir.

HSS'in üst seviye genel değerlendirilmesi iki şekilde yapıldı:

- ZeDeS uzman sistemi kullanılarak vurma_Oranı, reaksiyon_zamanı ve hasar_Oranı'nın bir fonksiyonu olarak
- ZeDeS bulanık mantık yazılımı kullanılarak, üyelik fonksiyonları ve bulanık mantık kurallarının işletilmesiyle

HSS için kullanılan bulanık mantık kuralları Tablo 2 (eğer vurma Oranı (hit ratio) iyiye (good)) ve Tablo 3'te (eğer vurma oranı kötüye (bad)) verilmiştir. Vurma oranı, reaksiyon zamanı (reaction time) ve hasarlı alan (destroyed area) girişler olarak ve genel performans çıkış olarak kullanılmıştır. Mesela, Tablo 2'deki altı çizili hücredeki kuralın ifadesi aşağıda verilmiştir:



Şekil 5. Hasar alanın hesaplanması

Eğer (reaksiyon zamanı ortaysa (average)) ve (vurma oranı iyiye (good)) ve (Hasarlı alan azsa (little)),
öyleyse (genel performans çok iyidir (very good)).

Diğer kurallarda, örnekte ifade edilen kurala benzer tarzda tanımlanmıştır.

Tablo 2. HSS'deki bulanık kurallar1((eğer vurma oranı iyiye (good))

Reaksiyon Zamanı	Bad (kötü)	Average (orta)	Vgood (çok iyi)
Hasarlı Alan			
Little (az)	excellent	<u>vgood</u>	vgood
Average (orta)	good	good	good
Large (büyük)	lgood	bad	bad

Tablo 3. HSS'deki bulanık kurallar2 (eğer vurma oranı kötüye (bad))

Reaksiyon Zamanı	Vgood (çok iyi)	Average (orta)	Bad (kötü)
Hasarlı Alan			
Little (az)	vgood	vgood	good
Average (orta)	lgood	bad	bad
Large (büyük)	bad	vbad	vbad

Reaksiyon zamanı (reaction time) ve Hasarlı alan (destroyed area) ve genel performans üyelik fonksiyonları, Şekil 7, Şekil 8 ve Şekil 9'da verilmiştir.

OBJ	OBJ_TITLE	RESULT	RESULT_V
44	ADSystemReactionTime	GOOD	5880
45	HitRatio	Sufficient	100
46	DestroyedShelteredArea	Damage on Defenced Ar	0.38
51	ADSystemOverallPerforma	ADSystem Overall Perfor	79.89

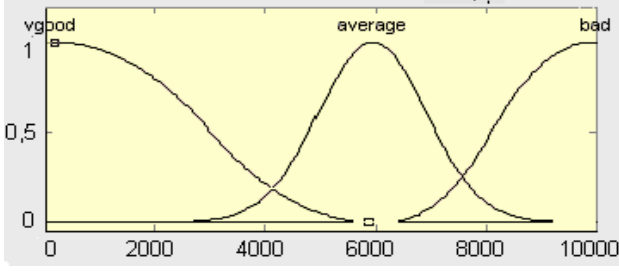
Reason

Used Rules:
Sheltered_Area_Ratio <10 " Damage on Defenced Area is Acceptable "

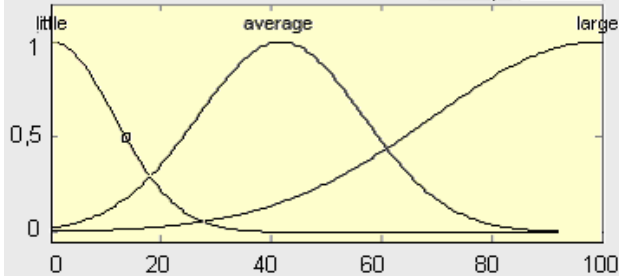
Used Methods:
Sheltered_Area_Ratio_Met=alfa:=ArcCos((Run^2+R^2-Rcm^2)/(2*Run*R)); beta:=ArcCos((Rcm^2+R^2-Run^2)/(2*Rcm*R));
Sheltered_Area_Ratio:=(100/3.14)*((Eval(alfa)-sin(2*Eval(alfa))/2)+(Rcm/Run)^2*(Eval(beta)-sin(2*Eval(beta))/2)); =0.384686545701;

Used Parameters:
R=1450 " Range " Run=1000 " UN camp radius "
Rcm=500 " Cruise Missile Destruction Radius "

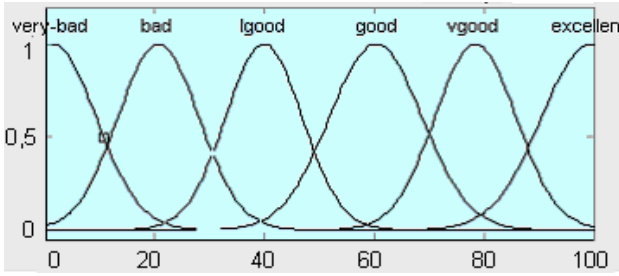
Şekil 6. ZeDeS'in HSS değerlendirme sonuç ekran görüntüsü



Şekil 7. Reaksiyon zamanı üyelik fonksiyonu



Şekil 8. Hasarlı alan üyelik fonksiyonu



Şekil 9. Performans üyelik fonksiyonu

ZeDeS bulanık mantık sisteminin genel performans sonuçları Şekil 10'da gösterilmiştir. Bu şekilde, bulanık çıkarım süreçleri paralel işler halinde gösterilmiştir. Sistemde toplam 18 kural olup her bir satır bir kuralın işleyişinin göstermektedir. Örneğin, yukarıda verilen Tablo 2'deki altı çizili hücredeki kuralın işleyişi 11 nolu satırda verilmiştir. Sağ altta gösterilen dikdörtgen, tüm kuralların çıkışlarının toplamının ağırlık merkezidir. Bu makalede incelenen literatürdeki değerlendirme sistemleri şunlardır:

- SIMULTAAN PASS Yazılımı (Arend ve Jansen, 2000)
- RTP 11.12 Projesi PES (Performance Evaluation System) Yazılımı (Öztemel vd., 2003a)
- RTP 11.13 Projesi EDST (Evaluation Definition Selection Tool), EDT (Evaluation

Definition Tool) ve EET (Evaluation Execution Tool) Yazılımları (Öztemel vd., 2003a; Rollesbroich vd., 2003b)

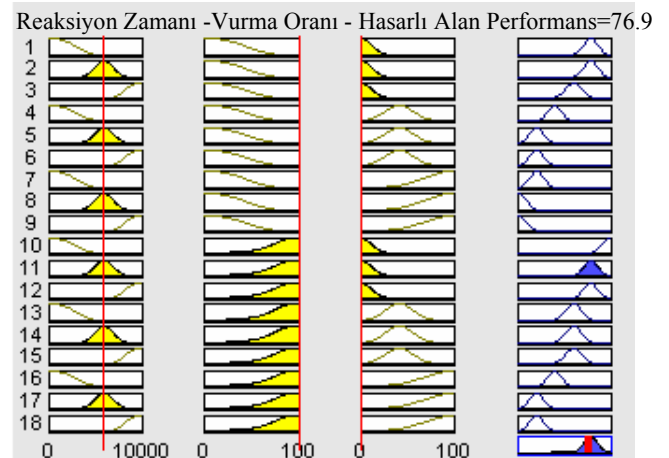
- ESTA (Expert System for Text Animation) Evaluation

Tablo 4'de ZeDeS'in benzer sistemlerle karşılaştırılması verilmiştir. ZeDeS'in benzer sistemlere göre bazı üstünlükleri şunlardır:

- ESSE (Expert System for Software Evaluation) (Vlahavas vd.,1998)
- ZeDeS, birbirinden çok farklı alanlarda kullanılabilirken, diğer sistemler ise belli bir alana yönelik olarak geliştirmişlerdir.
- ZeDeS ve ESTA, sonuçlarının nedeni açıklarken, diğerleri açıklayamamaktadır.
- ZeDeS, değerlendirme için gerekli olan bilgilerin yazılımın kaynak kodunun değiştirilmeden güncellenmesine olanak sağlar.

Sonuçlar

Bu çalışmada, değerlendirme sürecini kolaylaştıran ve hızlandıran Genel Değerlendirme Modeli (GeDeM) geliştirildi. Bu kapsamda, farklı alanlardaki uzmanlardan elde edilen sezgisel bilgilerin ve farklı kaynaklardan elde edilen verilerin değerlendirme amaçlı kullanılabilmesi için bir yöntem geliştirildi. Bu yöntemde, değerlendirme bilgileri, değerlendirme amaçlarının, değerlendirme kurallarının, ölçümlerin, metodların ve parametrelerin referans modeli olarak ifade edildi. Değerlendirme bilgilerinin referans modeli olan GeDeM, değerlendirme bilgilerinin



Şekil 10. Genel değerlendirme sonuçları

Tablo 4. ZeDeS'in benzer sistemlerle karşılaştırılması

	ZeDeS	EDST+ EDT+ EET	SIMULTA AN PASS	PES	ESTA Evaluation	ESSE
Yapay Zeka Metodolojisi	Tümleşik Uzman Sistem ve Bulanık Mantık	Kısmen Uzman Sistem	Kısmen işlenebilen kurallar	Yok	Uzman Sistem	Uzman Sistem
Uygulama Alanı	Genel (Sistem, sentetik ortamlar, insan, vs)	Eğitim, tatbikat ve tedarik	Eğitim	Pilot Değerlendirme	Sentetik Ortamlar	Yazılım
Temel Amacı	Sistem, sentetik ortamlar ve insanların değerlendirilmesi	Sentetik ortamların değerlendirilmesi	Eğitilen kişilerin ve takımların performans değerlendirilmesi	Pilot Performans Değerlendirme	Sentetik ortamların değerlendirilmesi	Yazılım problem çözme ve yazılım özellik değerlendirme
Bilgi Detayları	Değerlendirme amaçları, kuralları, ölçütler, metotlar, parametreler	Değerlendirme amaçları, kuralları, ölçütler, metotlar, parametreler	Senaryoya özel eylemler ve ilgili kurallar	Pilot Değerlendirme amaçları, indeksler, kuralları	Değerlendirme Bilgi Tabanı	Çoklu kriterler, yazılım öznitelikleri
Bilgi Gösterimi	Kural Tabanlı	Kural Tabanlı	Eylemle ilişkili değerlendirme kuralları	Algoritmik	Kural Tabanlı	Kural Tabanlı
Bilginin kaynak kodu değiştirilmeden güncellenebilmesi	Evet	Kısmen	Kısmen	Yok	Evet	Kısmen
Sonuca nasıl ulaşıldığının açıklanması	var	yok	yok	yok	var	Kısmen var
Kullanılabilirlik	Kolay	Zor		Kolay	Zor	

ifade edilmesi için gerekli kuralların, ölçümlerin ve parametrelerin sayısını azalttı. Zeki Değerlendirme Sistemi (ZeDeS), GeDeM'e ve değerlendirme ihtiyaçlarına göre geliştirildi. Melez uzman sistem ve bulanık mantıktan meydana gelen ZeDeS yazılımının değerlendirme amacıyla kullanılabileceği ve değerlendirme sürecini hızlandırabileceği gösterildi. ZeDeS'in sağlayacağı bazı faydalar kısaca şu şekildedir:

- Değerlendirme zamanını kısaltır.
- Değerlendirme bilgilerinin modellenmesinde ve işlenmesinde esnek bir yapı sağlar
- Farklı alanlarda kullanılabilir
- Değerlendirme sonuçlarına nasıl ulaşıldığını açıklar.
- Değerlendirme sürecinde kullanılan bilgilerin paylaşılmasını ve tekrar kullanılmasını sağlar.

- Değerlendirme için gerekli olan bilgilerin kaynak kodu değiştirilmeden güncellenmesine olanak sağlar.
- Değerlendirmeye ilgili karmaşıklığı azaltır.
- Üst seviye değerlendirmeyle ilgili olarak belirsizliklerin modellenmesinde ve sözel ifadelerle çıkarımda kullanılabilir.

Değerlendirme bazı yönleriyle belirsizlik içerdiğinden ve uzmanlara göre farklılık gösterdiğinden dolayı bulanık mantık, ZeDeS'in çıkarım motorunda karar vermede (reasoning) kullanıldı. Fakat bulanık mantığın değerlendirmenin tüm kısımları yerine, üst seviye genel (özet) değerlendirmede daha kolay uygulanabileceği ve daha faydalı olabileceği görüldü. Bunun nedeni ise, komple bir değerlendirmede birçok parametrenin gerekmesi ve bulanık mantıkta tüm bu

parametrelerle ilgili olarak birçok kural yazmanın etkili ve kolay bir yol olmayacağındandır. Sistemler karmaşıklıkça birçok kural gerektirir ve sistemle bu kurallar arasında ilgi kurmanın zorluğu geometrik olarak artacaktır. Bu kural sayısı yaklaşık olarak 15 geçtiği zaman ilgi kurma kabiliyeti tipik olarak azalacaktır (Lakhmi ve Martin, 1998). Bunlardan dolayı bulanık mantık en üst seviye değerlendirme için kullanıldı.

Teşekkür

Çalışmada desteklerini esirgemeyen TÜBİTAK MAM Bilişim Teknolojileri Ens. ve RTP11.13 projesi yetkili ve çalışanları teşekkürle anılır.

Kaynaklar

- Bass, E. J., (1998). Architecture for an intelligent instructor pilot decision support system, *IEEE International Conference on Systems, Man, and Cybernetics*, 1, 891–896, San Diego, CA.
- Drewes, P. ve Gonzalez, A., (1994). Automatic performance monitoring and evaluation, *IEEE Proceedings of the Fifth Annual Conference on AI, Simulation, and Planning in High Autonomy Systems, Distributed Interactive Simulation Environments*, 274-280, Gainesville, FL, USA.
- Gregory W. H., (1998). Evaluating Simulation Based Training Scenarios, *1998 Spring SIW (Simulation Interoperability W.)*, Orlando, FL.
- Hartley R. ve Varley G., (2001). The design and evaluation for development of complex decision making skills, *Proc. IEEE International Conference on Advanced Learning*, Madison, WI, USA.
- Konar, A., (2000). Artificial intelligence and soft computing behavioral and cognitive modeling of the human brain, CRC Press, Florida, USA.
- Lakhmi C. J. ve Martin N.M., (1998). *Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications*, CRC Press, Florida, USA.
- Oliveira J.C., Shirmohammad S., ve Georganas N.D., (1999). Collaborative Virtual Environment standards: A performance evaluation, *Proc. 3rd IEEE Inter. Workshop on Distributed Interactive Simulation and Real-Time Applications*, 14–21 Oct., 22-23, Greenbelt, MD, USA.
- Öztemel E., Gürbüz A. ve Görçin A., (2003a). In-flight demonstration of embedded simulation for training purposes on-board fighter aircraft, Technical Report, TÜBİTAK-MAM, Turkey.

- Öztemel E., Öztürk V., Soyer B. S., Öztürk S. ve Gonidakis V. (2003b). Inference engines specifications, Technical Document, RTP 11.13 Project Technical Report, TÜBİTAK-MAM.
- Öztemel E., ve Öztürk V., (2003). Intelligent evaluation definition of training systems in synthetic environments, *ITEC Conference*, 29 April -1 May 2003, London, UK.
- Rigg, G., Morley R. ve Hepplewhite R., (2000). Themis: The objective assessment of CGF performance, *SISO 9th Conf. on Computer Generated Forces & Behavioral Representation*, 16-18 May 2000, Orlando, FL.
- Rollesbroich, B., zu Drewes, P. M., Greiwe, K. Jokipii, M. ve Hartikainen, T., (2003a). Common evaluation framework & evaluation knowledge editor, RTP11.13 –CAE–WE6.1–TR–1.0, *RTP 11.13 Project Technical Report, CAE, Germany*.
- Rollesbroich B., Öztürk V., Öztemel E., Soyer B. S., Öztürk S. ve Gonidakis V., (2003b). Knowledge base description, Technical Document, RTP 11.13 Project Technical Report, *CAE, Germany*.
- Russell, S. ve Norvig, P., (1995). Artificial intelligence: A modern approach, Prentice Hall, New Jersey, USA.
- Shub Y., Kushnir A., Frenkel J., (1994). Pilot evaluation system, *Proc. of IEEE Aerospace and Electronics Conf.*, 734–741, 23-27 May 1994, Dayton, OH, USA.
- Vlahavas, I., Stamelos, I., Refanidis, I. ve Tsoukias, A., (1998). ESSE: an expert system for software evaluation, *Knowledge-Based Systems*, 183–197.
-
- Arend J. ve Jansen R., (2000). Individual and team performance assessment in networked simulation training, <http://www.nlr.nl/public/hosted-sites/simultaan/simultaan.pdf>, (12.09.04)
- Hornung B., (1995). “Bernd H.’s Glossary”, <http://pespmc1.vub.ac.be/ASC/hornung.html> (05.11.04)
- Macannuco, D., Hung, J. ve Civinskas, W., (1998). A Test Suite to Evaluate Run-Time Infrastructure (RTI) Implementations for High Performance, Human-In-The-Loop (HITL) Simulators, http://www.sisostds.org/doclib/doclib.cfm?SISO_FID_1260 (05.11.2004)
- Tidhar, G., Heinze, C., Goss, S., Murray, G., Appl, D. ve Lloyd, I., (1999). American Association for Artificial Intelligence, www.aaai.org
- Trochin W. M. K., (2004). Introduction to Evaluation, <http://www.socialresearchmethods.net/kb/intreval.htm> (05.11.2004).