

ÖZET

TECRÜBE İLE SÜREKLİ ÖĞRENEN OTONOM SANAL SİMÜLATÖR VARLIKLARI

Buluş, eğitim simülasyonlarında (3) kullanılan sanal varlıkların (2), kural tabanlı
5 davranış yapısından yapay zekâ tabanlı öğrenen davranış yapısına geçirilmesinin
sağlandığı ve bu sayede sanal varlıkların (2) kendini geliştirdiği bir yapay zekâ
modeli tasarlanması ve bu sayede sanal varlıkların (2) kendini geliştirdiği bir
yöntem ile ilgilidir. Simülatörde eğitim algoritması (1) olarak denetimli öğrenme
ve pekiştirmeli öğrenme algoritmaları ile geliştirilen ve simülasyona (3)
10 uygulanarak öğrenen otonom sanal varlıklar (2) tasarlanmaktadır.

İSTEMLER

1. Buluş eğitim simülatörlerinde bir yapay zekâ modeli ile sanal varlıkların kendini geliştirme yöntemi olup,
 - Denetimli öğrenme eğitim algoritmaları (1) ile yönetilen akıllı sanal varlıklar (2) için derin öğrenme algoritması ile optimum bir davranış modeli (5) geliştirilmesi,
 - Eğitim algoritmalarındaki (1) eğitim arayüzüne (4) simülasyonların (3) akıllı sanal varlıkla (2) bağlanması ve bu sayede akıllı sanal varlıkların (2) bir arayüz gibi davranması,
 - Denetimli öğrenme eğitimin ardından pekiştirmeli öğrenme ile ödül ceza veya amaç fonksiyonuna göre deneme yanılma ile kendisi için uygun olan davranış modelini (5) öğrenmesi,
- adımlarını içermekte ve
- Akıllı sanal varlıkların (2) farklı konumlardan başlatılarak simülatör ortamını genellemesi ve
 - Akıllı sanal varlıkların (2) eş zamanlı olarak senaryolar arası genelleme yaparak paralel çoklu eğitim yapabilme kabiliyetine sahip olması ile karakterize edilmektedir.
2. Buluş istem 1'deki gibi bir yöntem olup, eğitim algoritmalarının (1) derin öğrenme algoritmaları olması ile karakterize edilmektedir.

TARİFNAME

TECRÜBE İLE SÜREKLİ ÖĞRENEN OTONOM SANAL SİMÜLATÖR VARLIKLARI

Teknik Alan

- 5 Buluş, eğitim simülasyonlarında kullanılan sanal varlıkların, kural tabanlı davranış yapısından yapay zekâ tabanlı öğrenen davranış yapısına geçirilmesinin sağlandığı ve bu sayede sanal varlıkların kendini geliştirdiği bir yapay zekâ modeli tasarlanması ve bu sayede sanal varlıkların kendini geliştirdiği bir yöntem ile ilgilidir.

10 Önceki Teknik

Simülasyonlardaki kural tabanlı varlık kontrol sistemleri gerçek bir sistemi modellemek için yetersiz kalmaktadır ve kuralların hazırlanması özellikle çok fazla varlığın olduğu bir ortamda ciddi miktarda vakit almaktadır.

- 15 Eğitim amaçlı kullanılan simülasyonlarda eğitim alan kişilerin, kural tabanlı çalışan bu sistemlerde kendilerini geliştirme seviyesi doğrudan simülasyonları hazırlayan kişilere ve kurallara bağlı olduğu için canlı eğitimlerin oldukça gerisinde kalmaktadır.

- 20 Bu kural bazlı simülasyon yöntemleri eğitim alan kişileri tatmin etmemekle beraber, risk kapsamı yüksek işlerde gelecek vadetmemektedir. Canlı eğitimlerde oldukça fazla olası kombinasyonların var olması, ortaya çıkabilecek sorunlar ve bunlara çözüm üretme ihtiyacı, eğitimin kalitesini artıran faktörlerdendir.

Simülasyonda eğitim alan kişi ve kural bazlı sistem arasında kendini yenileyen bir sistem olmadığı için sabit yöntemler ile ilerleyen bu sanal ortamlar yeni metotlar

geliştirerek tecrübe kazanma idealine oldukça uzaktır, dolayısıyla simülasyonlar geleneksel yöntemlerle, gerçek harekât ortamı seviyesine yaklaşamamaktır.

Eğitim simülasyon programlarında, ortamdaki sanal varlıklar insan kontrolü veya önceden programlanmış kural tabanlı alışlagelmiş yöntemlerle kontrol edilir.

- 5 Eğitim simülasyon programlarında kursiyerlerin sanal ortamda olabildiğince farklı durumlar içinde kalması ve bu sayede tecrübe kazanması beklenir. Kural tabanlı davranışların en önemli dezavantajlarından biri bu belli ve sabit senaryo şartlarına uyumlu olarak önceden programlanmasıdır. Eğitim sırasında gerçekleşen olası bir varyantla karşılaşıldığı zaman kural tabanlı sanal varlıklar etkisiz kalmaktadır.
- 10 Yapay zekâ tabanlı davranış modelleri kullanan sanal varlıklar genelleme yaparak eğitildiği için bu problemin üstesinden gelebilmektedir. Spesifik durumlara özgü modellenmeyen, beklenmedik durumlara karşı çözüm üretebilen bu varlıklar her zaman için geliştirilebilir ve ortama adaptasyon sağlayabilen bir nitelik taşımaktadır.
- 15 Kompleks simülasyon ortamlarında çok sayıda sanal varlığın etkileşimi söz konusu olabilmektedir. Ortamdaki kompleksite arttıkça bu varlıkların davranış mekanizmalarının programlanması geleneksel yaklaşımlarla mümkün değildir. Yapay zekâ yaklaşımları ile ortamdaki varlıklar ve unsurlar arasındaki etkileşimler öğrenilebileceği için ortamdaki varlık sayısından bağımsız olarak sistem
- 20 sürdürülebilirliği sağlanabilmektedir. Bu durum hem kural yazarı (alan uzmanı ve yazılımcı) hem kullanıcıyı önemli bir problemten kurtarmaktadır.

Bir eğitim simülatörlerinden beklenen en önemli özelliklerden biri de gerçekçiliktir.

Katı ve durağan kural tabanlı kontrol mekanizmaları sanal varlıklara robotik karakterler verdiği için simülasyon ortamının gerçekçiliği zedelenmektedir. İnsana

- 25 benzer olarak zaman içinde tecrübelerinden öğrenen, kendini geliştirebilen, içinde bulunduğu duruma göre farklı stratejiler uygulayabilen akıllı sanal varlıklar simülasyon gerçekçiliğine olumlu katkı sağlamaktadır.

Eğitim esnasında eğitim simülatörlerinden maksimum verim alabilmek için mücadele edilen varlıkların davranış modellerinin, eğitim görecekt kursiyerlerin tecrübe seviyesine göre belirlenmesi yaygın kullanılan bir yöntemdir. Fakat bu durum kural tabanlı davranış modelleri hazırlanırken çok çeşitli varyasyona uyum sağlayabilme ihtiyacı anlamına gelmektedir. Dolayısıyla farklı tecrübe seviyelerine uyumlu kural tabanlı davranış modelleri tasarlama durumu söz konusudur.

- Akıllı sanal varlıklar ise kendilerine rakip olacak kursiyerlerin seviyesinde eğitilerek bu eşleşme sağlanabilir. Açıkçası sanal varlıkların davranışlarının gerçekçi modellenmesi, eğitim alanların doğru yönde geliştirilmesi için önemlidir.
- 10 Mevcut geleneksel yaklaşımlarla eğitim gören kursiyerler, kendilerini kural tabanlı davranış mekanizmalarını tasarlayan uzmanların ve yazılımcıların ilgili alandaki tecrübe seviyesi ölçüsünde geliştirebilirler. Bu açıdan bakıldığında ilk aşamalarında temel bilgilerini konunun uzmanlarından öğrenen yapay zekâ algoritmaları kendini geliştirdikçe insanlara yeni çözüm önerileri sunabilir.
- 15 Tekniğin bilinen durumunda yer alan US2020293862A1 sayılı Birleşik Devletler Patent dokümanında, bir ortamla etkileşime giren bir aracı tarafından gerçekleştirilecek eylemleri seçen bir veya daha fazla konumdaki bir veya daha fazla bilgisayarda bilgisayar programları olarak uygulanan bir takviye öğrenme sistemini tanımlar. Genel olarak, sistem, çevre durumlarını karakterize eden gözlemlere yanıt olarak gerçekleştirilecek eylemleri seçerken bir eylem seçim politikası sinir ağı kullanır.
- 20

- Tekniğin bilinen durumunda yer alan CN106952014A sayılı Çin Patent dokümanında, bir askeri simülatöre dayalı bir savaş planı optimizasyon yönteminden bahsedilmektedir. Yöntem aracılığıyla, bir savaş planı otomatik olarak oluşturulabilir ve optimize edilebilir. Bu nedenle, askeri karar verme personeli için daha iyi bir askeri plan yardımcı karar verme işlevi sağlanır.
- 25

Tekniğin bilinen durumunda yer alan CN106682351 sayılı Çin Patent dokümanında, savaş alanı simülasyon eğitiminin teknik alanı ile, özellikle bir savaş simülasyon sistemi ve bilgisayarda üretilen kuvvetlere dayalı simülasyon yöntemi ile ilgilidir. Buluş bir bilgisayara ve bir simülasyon yöntemine dayalı olarak askeri güç üreten bir dövüş simülasyon sistemi sağlar ve savaş alanlarında simülatif eğitimin teknik alanı ile ilgilidir. İnsan davranışları üzerinde yeterince modelleme yaparak, simülatif bir savaş alanı ortamında insan etkileşimi olmadan olaylara ve durumlara otomatik olarak yanıt verebilir ve gerçek dövüşlerdeki olası sonuçlar simüle edilebilir bir yöntemden bahsedilmektedir.

10 Faaliyet gösterdikçe öğrenen akıllı sanal varlıklar henüz ilgili alandaki uzmanlar veya literatür tarafından henüz keşfedilmemiş yeni stratejiler geliştirerek eğitim simülasyonunun gerçekçiliğini artırmakta ve kursiyelerin doğru yönde geliştirilmesi için bir potansiyel oluşturmaktadır. Görev Simülatörlerinde belirlenmiş bir senaryoya uygun olarak hareket etmesi gereken taktik unsurların 15 simülasyonuna ihtiyaç bulunmaktadır. Bu taktik unsurların tecrübelerden öğrenerek bir zekâyâ ve adaptasyon yeteneğine sahip olması önemlidir. Bu şekilde kendisine verilen rolleri gerçek dünyadakine benzer şekilde gerçekleştirebilecek ve taktik simülasyonun amaçlarına daha iyi hizmet etmesini sağlayabilecektir.

Bu nedenle bir eğitim simülasyonlarında sanal varlıkların kendini geliştirdiği bir yapay zekâ modeli geliştirilmesine ihtiyaç duyulmuştur.

Buluşun Amacı

Bu buluşun amacı, görev simülatörlerinde belirlenmiş bir senaryoya uygun olarak hareket etmesi gereken taktik unsurların içeren simülasyonun gerçekleştirilmesidir.

Buluşun başka bir amacı, taktik unsurların tecrübelerden öğrenerek bir zekâyâ ve 25 adaptasyon yeteneğine sahip simülasyonun gerçekleştirilmesidir.

Bu buluşun başla bir amacı, kendisine verilen rolleri gerçek dünyadakine benzer şekilde gerçekleştirebilecek ve taktik simülasyonun amaçlarına daha iyi hizmet etmesini sağlayabilecek bir eğitim simülasyonunun gerçekleştirilmesidir.

Buluşun Ayrıntılı Açıklaması

- 5 Bu buluşun amaçlarına ulaşmak için gerçekleştirilen eğitim simülatörlerinde bir yapay zekâ modeli ile sanal varlıkların kendini geliştirme yöntemi ekli şekilde gösterilmiştir.

Bu şekil;

- 10 **Şekil 1:** Buluş konusu sanal varlıkların kendini geliştirdiği bir yapay zekâ modelinde yer alan temel unsurlarının şematik görünümüdür.

Şekilde yer alan sistemler tek tek numaralandırılmış olup, bu numaraların karşılıkları aşağıda verilmiştir.

1. Eğitim algoritmaları
2. Sanal varlıklar
- 15 3. Simülasyon
4. Eğitim arayüzü
5. Davranış modeli
6. Ham aksiyon
7. İşlenmiş aksiyon
- 20 8. Ham veri
9. İşlenmiş veri, ödül

Buluş eğitim simülatörlerinde bir yapay zekâ modeli ile sanal varlıkların kendini geliştirme yöntemi olup,

- Denetimli öğrenme eğitim algoritmaları (1) ile yönetilen akıllı sanal varlıklar (2) için derin öğrenme algoritması ile optimum bir davranış modeli (5) geliştirilmesi,
- Eğitim algoritmaları (1) ile simülasyonların (3) akıllı sanal varlıkla (2) 5 bağlanması,
- Eğitimin ardından pekiştirmeli öğrenme ile ceza veya amaç fonksiyonuna göre deneme yanılma ile kendisi için uygun olan davranış modelini (5) öğrenmesi,

adımlarını içermektedir.

- 10 Buluş konusu yöntem, 3 temel bileşenden oluşmaktadır. Bu yapı Şekil 1 üzerinde gösterilmiştir. İlk bileşen, eğitim algoritmaları (1) olup sistemin temelini oluşturmaktadır. Burada yönetilen akıllı sanal varlıklar (2) için optimum bir davranış modeli (5) geliştirilmeye çalışılmaktadır. Geliştirilen eğitim algoritmaları (1) derin öğrenme algoritmalarıdır. Eğitim yöntemleri ise denetimli öğrenme yani
- 15 supervised learning ve pekiştirmeli öğrenme yani reinforcement learning şeklinde ikiye ayrılabilir. Supervised learning bir kullanıcının yaptığı davranışlardan öğrenme olarak tanımlanabilir. Burada bir kullanıcı verisi var ise sistem bu veri ile eğitilerek belirli bir kapasiteye ulaştırılır. Sonrasında bu eğitimin üzerine pekiştirmeli öğrenme ile devam edilir. Bu eğitim sistemin deneme yanılma ile
- 20 öğrenmesi olarak tanımlanabilir. Sistem tasarlanan ödül ceza veya amaç fonksiyonuna göre deneme yanılma ile kendisi için uygun olan davranış modelini (5) öğrenmeye çalışır. Bu noktada sistem sıfırdan direkt olarak pekiştirmeli öğrenme ile de başlatılabilir ama bu durumda eğitim süresi çok daha uzun sürecektir.
- 25 İkinci bileşen olarak akıllı sanal varlıklar (2) bulunmaktadır. Bunlar, aslında simülasyonda (3) davranış gösteren varlıklardır. Dolayısıyla geliştirilen simülasyonda (3) varlıklara bir davranış modeli (5) atanabildiği varsayılır. Burada yapılan geliştirim, varlığı eğitim algoritmasına (1) bağlayarak bir arayüz gibi

kullanma tekniğidir. Sanal varlık (2) bu teknik sayesinde yapay zekâ ile geliştirilen davranış modelini (5) uygulayabilir hale gelir. Bu yapı eğitim algoritmalarında (1) “gym” olarak ifade edilen eğitim arayüzüne (4) bağlanır. Bu eğitim arayüzü (4) bir fonksiyon ile varlığın yapacağı aksiyonları, ham aksiyon (6), iletip bu aksiyon sonucunda elde ettiği yeni durum uzayını alır.

Simülasyon (3) ise üzerinde geliştirme yapılan gerçek dünyayı modelleyen yüksek sadakatli bir yapıdır. Burada geliştirilme yapılmamakta, simülasyona (3) işlenmiş aksiyonlar (7) iletilip, simülasyondan (3) gelen durum uzayı ile kararlar verilmektedir.

- 10 Buluş, simülasyon (3) ortamında önceliklendirilsede, simülasyonun (3) gerçek dünyayı ne kadar iyi modellediği ile orantılı olarak simülasyon (3) ortamının bulunduğu domain ile uyumlu yerlerde karar destek sistemi olarak da kullanılabilir. Çünkü sanal varlık (2) aslında karşısında bulunduğu koşullarda tasarlanılan ödül ceza sistemine göre en uygun davranışı yapmayı öğrenmektedir. Dolayısıyla benzer
- 15 koşullarda gerçek platformlar ve hareket ortamları için karar destek sistemi olarak tavsiyeler verebilir. Bunun yanında buluş, yapay zekâ sistemleri ile yüksek hesaplama kabiliyetleri kullandığından bir kullanıcının verebileceği karar hızından çok daha hızlı çalışabilmektedir. Dolayısıyla hızlı karar verilmesi gereken durumlarda oldukça faydalı bir sistem haline gelebilir.
- 20 Bunun yanında yapı kullanıcının seviyesine göre eğitim şeklinde bir duruma da evrilebilir. Bu sanal varlıkların (2) sürekli öğrenen bir yapı içinde olduğu düşünüldüğünde olası bir çalışmada kullanıcının seviye seçmesine olanak sağlayacak bir yapıya evrilebileceği düşünülmektedir. Örneğin az miktarda eğitilmiş sistem kolay seviye olarak kullanıcının karşısına çıkarken daha fazla
- 25 eğitilmiş sistem daha profesyonel bir seviye olarak eğitim alan kullanıcıya sunulabilir.

Kural tabanlı çalışan sanal ortam simülasyonları (3) için hazırlanan senaryoların kapsamı ve gerçekçiliği senaryoyu hazırlayan kural yazarlarının yetisiyle sınırlıdır. Bu sınırlamayı ortadan kaldırabilmek amacıyla, kendi kendine öğrenen ve bilmediği durumlarla karşılaştığında daha doğru kararlar alabilen sanal varlıkların (2) simülasyon (3) ortamına eklenmesi düşünülmüştür.

Temel olarak problemi çözmek için çalışma basamaklarını sıralarsak 8 basamak belirtilebilir. Bunlar durum uzayı analizi, aksiyon uzayı analizi, simülasyon (SDK) analizi, gym (eğitim arayüzü) arayüzünün (4) hazırlanması, varsa kural tabanlı davranış modellerinin veya uzman verilerinin incelenmesi, denetimli öğrenme algoritmalarının uygulanması, pekiştirmeli öğrenme algoritmalarının uygulanması ve tekli/çoklu varlık eğitimleri olarak sıralanabilir.

- **Durum Uzayı Analizi**

Durum uzayında simülasyondan alınan Tablo-1' deki veriler kullanılmaktadır:

Tablo-1: Durum Uzayı Verileri

Durum Uzayı Verileri	
1	Görev seçimi yapılan sanal varlığın ve muhtemel muhasımlarının konum, yükseklik, hız, yön, oryantasyon öznitelikleri
2	Görev seçimi yapılan sanal varlığın ve muhtemel muhasımlarının sınıf ve tip öznitelikleri
3	Görev seçimi yapılan sanal varlığın sensör kabiliyetleri ile elde ettiği durumsal farkındalığı
4	Görev seçimi yapılan sanal varlığın kalan mühimmatlarının tipini ve sayısı

5	Görev seçimi yapılan sanal varlığın IFF kullanımı sonucunda elde ettiği farkındalığı
6	Görev seçimi yapılan sanal varlığın en fazla bir dost birimle haberleşerek oluşturduğu taktiksel farkındalığı ve taktiksel görev atamaları
7	Görev seçimi yapılan sanal varlığın muhtemel muhasımının karşı tedbir manevra kabiliyetleri hakkında kestirimleri
8	Görev seçimi yapılan sanal varlığın sahip olduğu rotanın özellikleri
9	Hava şartları, günün zamanı ve arazi koşulları
10	Görev seçimi yapılan sanal varlığın radyo mesajı kullanarak elde ettiği durumsal farkındalığı ve taktiksel görev atamaları
11	Görev seçimi yapılan sanal varlığın hasar durumu
12	Görev seçimi yapılan sanal varlığın yakıt durumu
13	Görev seçimi yapılan sanal varlığın içinde bulunduğu formasyonun özellikleri

- **Aksiyon Uzayı Analizi**

Hedeflenen buluşta aksiyon analizinde kullanılacak algoritmanın temel eylemleri ve parametreleri Tablo-2 ve Tablo-3’de verilmiştir.

Tablo-2: Temel eylem kategorileri

Temel Eylem Kategorileri	
1	Manevra Eylemleri (Parametre detayları Tablo 2’de verilmiştir.)
2	Silah Eylemleri (Silah Tipi, Atış Modu)
3	IR Flare Eylemleri (Flare Sayısı)
4	Chaff Eylemleri (Chaff Sayısı)
5	IR Karıştırıcı Eylemleri (Aç, Kapa, Oto)
6	Radar Karıştırıcı Eylemleri (Aç, Kapa, Oto)
7	Lazer Karıştırıcı Eylemleri (Aç, Kapa, Oto)
8	IFF Eylemleri (Mod, Kod)
9	Duman Eylemleri (Aç, Kapa, Oto)
10	Sensör Eylemleri (Mod, Aç, Kapa)
11	İletişim Eylemleri (Mesaj Gönder, Mesaj Al)

Tablo-3: Temel eylemler ve parametreleri

Temel Eylemler	Parametreler
Takip Et	Takip edilecek platform kimlik bilgisi
	Takip minimum / maksimum mesafesi (metre)
	Takip açısı (derece)
	Takip irtifası (metre)
Konuma Git	Koordinat bilgisi

	Beklenen ulaşma zamanı veya İlerleme Hızı (knots)
	Dönüş Hızı (Opsiyonel) (derece/sn)
	Tırmanış/ Dalış Hızı (Opsiyonel)
	Hızlanma / Yavaşlama İvmesi (Opsiyonel) ()
	İrtifa (MSL/AGL) (Opsiyonel) (metre)
Oyuncuya Git	Oyuncu kimlik bilgisi
	Yaklaşma mesafesi (metre)
	Yaklaşma açısı (derece)
	Yaklaşma irtifası (metre)
	Dönüş Hızı (Opsiyonel)
	Tırmanış / Dalış Hızı (Opsiyonel)
	Hızlanma / Yavaşlama İvmesi (Opsiyonel) ()
Verilen Değerlerle İlerle Yön Bilgisi (Derece)	Yön bilgisi (derece)
	Hız bilgisi (knots)
	Dönüş Hızı (Opsiyonel)
	Tırmanış / Dalış Hızı (Opsiyonel)
	Hızlanma / Yavaşlama İvmesi (Opsiyonel) ()
	İrtifa bilgisi (metre)
Görev Rotasına Gir	Rota kimlik bilgisi
	Rota sapma mesafesi (Opsiyonel) (metre)
	Rota sapma açısı (Opsiyonel)(derece)
	Rota Takip Türü (Opsiyonel)

- **Simülasyon (SDK) Analizi**

Muharip platformların simülatörlerinde taktik eğitim verilmesi amacıyla kullanılan Taktik Çevre Yazılımları, taktik senaryo planlama, planlanan senaryonun koşulması, eğitim sonrası değerlendirme ve dibrifing imkânı sağlayan simülasyon yazılımlarıdır.

Mevcut durumda; özellikle muharip platformlara ait görev simülatörlerinde, silah, sensör, karşı tedbir, kendini koruma ve haberleşme sistemlerinin eğitiminin verilebilmesi için, kapsamlı taktik eğitimlerin verilebilmesi için olabildiğince gerçekçi bir sanal bir taktik harekât ortamına ihtiyaç duyulmaktadır. Bu sanal harekât ortamındaki muhtemel dost, düşman ve tarafsız kuvvetler rollerini gerçekleştirebilen, taarruz ve/veya savunma görevlerini gerçekçi bir şekilde icra edebilen varlıklara ihtiyaç duyulmaktadır.

Bu varlıklar, hem fiziksel (aerodinamik, menzil, tespit, optik vb.) olarak gerçekçi dünyayı yansıtabilmeli, hem de gerçek hayattaki operatörlerinin davranışlarını (doktrinler, taktikler, tecrübe vb.) gerçekçi bir şekilde bu sanal dünyaya yansıtabilmelidir.

- **“gym” (eğitim arayüzü) Arayüzünün (4) Hazırlanması**

Gym, pekiştirmeli öğrenme algoritmalarını geliştirmek ve karşılaştırmak için geliştirilmiş eğitim ortamıdır. Bu ortam içerisinde eğitilen sanal varlıkların arkaya aksiyon alabilmesi ve alınan aksiyona göre alacağı ödül (9) ve bu aksiyonun sebep olduğu ortam içerisindeki durum değişiklikleri verimli bir eğitim için olması gereken temel özelliklerdir. Sanal varlığın eğitilebilmesi ve aynı ortamın içinde sıralı aksiyonlarla hedefine ulaşarak maksimum ödülü (9) alabilmesi için aynı senaryoyu tekrar tekrar koşması gerekmektedir.

Mevcut buluşta ilgili simülasyon içerisindeki sanal varlıkların aksiyon alabilmesi ve simülasyon ortamından durum uzayı bilgisinin her adımda sanal varlığın

hafızasına aktarılabilmesi için gym eğitim arayüzü (4) hazırlanmıştır. Gym eğitim arayüzünün (4) temel fonksiyonları aşağıda bahsedildiği gibidir. Diğer fonksiyonlar simülasyon ile sanal varlığın thread yapısında veri aktarımında kullanılmaktadır.

- 5 - step: her adımda sanal varlığın seçtiği aksiyonları alarak, simülasyon içerisine thread yapısıyla gönderip, gelen cevaba göre yeni durum değerlerini, hedefe ulaşıp ulaşmama durumunu ve alınan aksiyona göre ödül ceza sisteminden gelen ödül (9) değerini döndüren fonksiyon
- reset: eğitilen sanal varlığın maksimum aksiyona ulaşması, ölmesi gibi durumlarda simülasyonu ilk haline getirmeyi sağlayan fonksiyon
- 10 - get_reward: sanal varlığın aldığı aksiyona göre ödül (9) miktarını hesaplayıp bu değeri step fonksiyonuna aktaran fonksiyon

Günümüzde yaygın kullanılan pekiştirmeli öğrenme algoritmaları geliştiriciler tarafından “gym” ortamlarını kullanabilecek şekilde tasarlanmıştır. Dolayısıyla pekiştirmeli öğrenme algoritmalarında kullanılan genel geçer standart bir ana yapı oluşmuştur. Buluş bünyesinde “gym” eğitim arayüzünün (4) kullanılması ise buluşun algoritmik yapısının herhangi bir pekiştirmeli öğrenme algoritmasına adapte olabilmesine olanak sağlamıştır. Diğer bir deyişle literatürdeki herhangi bir pekiştirmeli öğrenme algoritması buluş üzerinde kullanılabilirken buluş üzerinde hâlihazırda kullanılan eğitim algoritması standart açık kaynak “gym” ortamlarında kullanılabılır.

Gym eğitim arayüzünün de (4) kendine göre belirli alt çalışmaları vardır. Bunların en başında ödül ceza fonksiyonu gelmektedir. Durum ve aksiyon uzayı, gym içinde modellenmektedir.

- 25 • **Varsa Kural Tabanlı Davranış Modellerinin veya Uzman Verilerinin İncelenmesi**

Kural tabanlı davranış modellerinin incelenmesinin en temel amacı simülatör ortamında durum bilgisi gözlemlene ve aksiyon uygulama prosedürlerinin anlaşılmasıdır. Sanal varlığın bünyesinde gömülü olan kural tabanlı davranış mekanizması buradan çıkartılıp simülatör dışı ortama alınırsa buluş bünyesindeki durum bilgisi gözlemlene ve aksiyon uygulama trafiğinin doğruluğu gözlemlenebilir hale gelir. Eğer sanal varlık beklenen davranışları göstermiyorsa bu durum buluş bileşenlerinden “gym” eğitim arayüzünde (4) potansiyel bir probleme işaret ederek geliştiriciye yol gösterir. Dolayısıyla kural tabanlı davranış modeli (5), buluşun sistemsel doğruluğu için referans role sahiptir. Ayrıyeten buluşun eğitim algoritmasında kullanacağı durum uzayı için karar verme açısından mantıklı olabilecek verilerin belirlenmesini sağlar. Diğer yandan kural tabanlı davranış modelleri, pekiştirmeli öğrenmenin temel taşlarından biri olan ödül fonksiyonunun tasarımı için genellikle faydalı ipuçları bulundurmaktadır. Kural tabanlı uygulanan savunma ve saldırı stratejileri analiz edilerek ödül fonksiyonu bu doğrultuda şekil almaktadır. Ayrıca kural tabanlı davranış modelinin (5) her adımda gözlemlendiği durum bilgisi ve aldığı aksiyon kaydedilerek bir uzman veri seti elde edilir. Bu veri seti denetimli öğrenme algoritmalarında kullanılmaktadır. Eğitim algoritmasının değerlendirilmesi konusunda da söz konusu analizden yararlanılmaktadır. Ödül fonksiyonunun kural tabanlı davranış modeline (5) verdiği ortalama skor bir referans kriter olarak baz alınıp öğrenme algoritmasının yakaladığı ortalama skorun referans skora olan uzaklığından yararlanarak performans hakkında fikir sahibi olunabilir.

- **Denetimli Öğrenme Algoritmalarının Uygulanması**

Denetimli öğrenme algoritmalarında yapay zekâ modeline verilen verinin, neye ait olduğu bilgisi, etiket, veri içindeki her parçaya uygulanmaktadır. Bu sayede model, girdi ile çıktı arasında bir ilişki kurabilmektedir. Eğitim sonrasında modele hem eğitim verisi hem de aynı içerikte fakat eğitimde kullanılmamış test verisi verilerek; modelden bu verilerin verilen numune sayısı kadar çıktısını (etiket, label İng.)

tahmin etmesi beklenmektedir. Doğruluk ya da hata değerine göre başarılı bulunan modeller istenen işlevi yerine getirmiş demektir.

Mevcut buluşta bahsedilen sanal varlıkların pekiştirmeli öğrenme ile eğitilmesinin yanında, rota tahmini gibi ek görevlerde denetimli öğrenme kullanılmaktadır.

- 5 Burada eğitim için gerekli veri seti, kural tabanına göre belirli bir senaryoda koşturulan sanal varlık üzerinden toplanmaktadır. Toplanan veri girdi ve çıktı olarak ayrılmaktadır. Girdi verisi, ham veri (8), sanal varlığın sahip olduğu yakıt durumu, konum bilgisi gibi değerleri içermektedir. Çıktı verisi, işlenmiş veri (9), modelden beklenen amaca göre değişiklik göstermektedir. Örneğin sanal varlığın
- 10 her aksiyon aldığı anda gelecek aksiyonu almadan ne tarafa yöneleceğini tahmin etmek isteniyorsa, anlık girdi değerlerine göre sonraki aksiyon alındıktan sonraki rota bilgisi çıktı olarak ayrılmaktadır. Bu sayede model girdi ile çıktı arasındaki ilişkiyi çözerek, rota tahmini yapabilecek hale gelmektedir.

Denetimli öğrenme algoritmaları, pekiştirmeli öğrenme algoritmalarının eğitim süresini kısaltmak için de kullanılmaktadır. Uzman bilgisi dahilinde yazılan kural tabanı farklı senaryolarda koşturularak sanal varlıkların hangi durum değerlerinde hangi aksiyonları aldıkları, girdi ve çıktı verisi olarak toplanmakta, ardından hazırlanan model eğitilmektedir. Sanal varlık bu modeli hafızasına alarak tecrübelerine göre öğrenmeye devam etmektedir.

- 20 Denetimli öğrenme algoritmaları, pekiştirmeli öğrenmeye başlangıç noktası olarak kullanılabileceği gibi diğer çok kritik bir fonksiyonu daha vardır. O da sistemin durum ve aksiyon uzayının kontrolüdür. Burada bir davranış taklit edildiği için, eğer sistemde bir eksiklik var ise bu aşamada tespit edilmektedir.

- **Pekiştirmeli Öğrenme Algoritmalarının Uygulanması**

- 25 Pekiştirmeli öğrenme yapay zekanın deneme yanılmayla karşı tarafa karşı üstünlük sağlamasını sağlayacak yeni taktikleri keşfetmesini sağlayan öğrenme tekniğidir.

Temelinde insan öğrenmesine dayanan bu teknik, varlıklara herhangi bir ön bilgi verilmeksizin keşfederek öğrenme olanağı sağlar. Pekiştirmeli öğrenme tekniği kullanan varlık, alınan aksiyonlar sonucu ortamdan geri bildirim olarak alınan ödül ve ceza sistemleri ile kendini geliştirmeye çalışır. Bu teknik denetimli öğrenme ile bağlantılı olarak kullanılabilir. Simülasyondan toplanan veri setleri ile yapılan denetimli öğrenme eğitimlerinde başarı elde edildikten sonra, pekiştirmeli öğrenme metotları ile model geliştirilebilir.

Pekiştirmeli öğrenme kısmında temel olarak "Proximal Policy Optimization" (PPO) olarak ifade edilen algoritma kullanılmıştır. Bu algoritma önceki dönem sonu raporlarında da birkaç kez ifade edilen "actor" ve "critic" adı verilen yapılardan oluşmaktadır. Bu actor critic yapı bileşenleri literatürde yetkinliği ispatlamış birçok algoritma içerisinde kullanılmaktadır. "Actor" aynı zamanda "policy" olarak da ifade edilirken durum uzayını alıp aksiyonlara karar veren derin öğrenme yapısıdır. "Critic" ise yalnızca öğrenme sırasında kullanılmakta olup öğrenilen modeli test ederken kullanılmamaktadır. "Critic" de derin öğrenme yapısı olup eğitim sırasında ödülleri (9) tahmin etmek için kullanılır. Ödülleri (9) tahmin eden bu yapı bazı algoritmalarda Bellman denklemlerini sağlamak için, bazı algoritmalarda ise ödül (9) varyansını azaltmada kullanılmaktadır. Critic'in olmasının diğer bir avantajı ise episode sonuna gitmeden ara noktadaki deneyimlerden yapay zekânın güncellenmesine olanak sağlamasıdır. Bahsedilen algoritmalar ile oluşturulan agentlar simülasyon içerisindeki sanal varlıklara karşılık gelmektedir. Bu sayede farklı senaryolarda gym eğitim arayüzü (4) ile pekiştirmeli öğrenme temelli sanal varlıklar eğitilmektedir. Eğitim yapılan senaryolar hava-hava, hava-yer olmak üzere, karada (Tank, S400 radar sistemi) ya da havada (SU27, F16 uçağı vb.) konuşlanmış sanal varlıkların birbirleri ile etkileşiminden oluşmaktadır.

- Üç Hava-Hava Senaryosu RL Eğitimi

Bu senaryo içerisinde eğitilen sanal varlıklar düşman sanal varlıklara üstünlük sağlama amacındadır. Senaryo içerisinde komuta kontrol sanal varlığı ve 2 yardımcı sanal varlık ile düşman tarafına ait 3 sanal varlık birbirlerine üstünlük sağlamaya çalışmaktadır. Beklenen çıktıda eğitilen sanal varlıklar düşman sanal varlıklara üstün gelmeli, doğru aksiyonlarla bu sanal varlıkları senaryo içerisinde pasif hale getirmelidir. Yapılan eğitimlerle düşman kuvvetlere ait sanal varlıklar, giderek öğrenen ve yeni stratejiler keşfeden sanal varlıklar tarafından alt edilmektedir.

- **Tekli/Çoklu Varlık Eğitimleri**

Yaygın kullanılan pekiştirmeli öğrenme ortamlarının çoğunda sanal varlığın sabit bir konumdan veya durumdan başladığı görülür. Dolayısıyla ortamların küçüklüğü ve kompleksitenin azlığı sebebiyle genelleme olgusu göz ardı edilir. Buluşta kullanılan simülatör ortamı gibi karmaşıklığın daha yüksek olduğu ortamlarda ise genelleme daha önemli bir problem haline gelir. Genellemenin konum, rol, ortamdaki platformlar gibi farklı kategorileri vardır. Akıllı sanal varlık farklı konumlardan başlatılarak konum açısından simülatör ortamını genellemesi ile bu problemin aşılması amaçlanmaktadır.

Diğer yandan buluşun paralel çoklu eğitim yapabilme kabiliyeti sayesinde akıllı sanal varlıklar birbirinden izole edilmiş kendi simülatör ortamlarında eğitim görebilmektedirler. Simülatör ortamlarına istenilen eğitim senaryosu yüklenebildiği varsayılırsa eş zamanlı olarak senaryolar arası genelleme yapmak da mümkündür. Buluş ile sağlanacak olan standart bir pekiştirmeli öğrenme eğitiminin literatürdekine benzer şekilde bazı dinamikleri vardır. Akıllı sanal varlığın simülatör ortamı içerisinde gözlem yapıp aksiyon uyguladığı birim zamana adım denilmektedir. Adım sayısı belli bir limite ulaştığında veya sonlandırma kriteri sağlandığında simülatör ortamının o anki durumu baştan başlatılır ve yeni bölüme geçilmiş olur. Akıllı sanal varlık söz konusu bölümler boyunca simülatör ortamında o anki duruma göre en mantıklı aksiyonları alarak optimal bir davranış modeline (5) yakınsamaya çalışır. Fakat benimsenen davranış modeli (5) her zaman en iyi

- özüm olmayabilir. Akıllı sanal varlık daha iyi stratejilerin varlığı ihtimaliyle bazen de diğ er aksiyonları deneyerek ortamı keşfetmeye alışır. Böylece mevcut davranış modelinden (5) daha optimal alternatif bir davranış modeli (5) olup olmadığını keşfetmiş olur. Buluş bünyesindeki loglama araçlarından eğitim metriklerinin durumu online olarak takip edilebilmektedir. Standart bir eğitim, akıllı sanal varlığın mantıksız ve rastgele aldığı aksiyonlar sonucu düşük bir skor ile başlar ve simülatör ortamının bölümler boyunca gözlemlenip yapılan keşifler sonucu skorun artması ve giderek daha mantıklı kararlar vererek maksimum ödüle (9) ulaşan bir davranış modelinin (5) öğrenilmesiyle sona erer.
- 10 Bunlara ek olarak analizleri gerçekleştirmek adına görsel araçlar da tasarlanmaktadır. Eğitim aşamasında anlık olarak çevrimiçi bir şekilde metrikleri loglayıp sonrasında metrikler gözlemlenmektedir. Ödül fonksiyonları kendi içinde de gruplandırıldığında bunlar da ayrı ayrı sistemde kaydedilebilmektedir. Eğitim durdurulup sonra kaldığı yerden devam ettirme özelliğine de sahiptir.
- 15 Buna ek olarak sistemimiz eğer kullanılan sistem imkân veriyorsa paralelden farklı eğitimleri yapabilmektedir. Bu sayede simülasyonun (3) imkân verdiği ölçüde farklı konseptteki eğitimleri paralelde alıştırarak karşılaştırma yapılabilir.

Şekil 1

1/1

