

SERVİS ROBOTLARI İÇİN ETİK BİRİMİ TASARIMI

K.Yasemin Usta ve H.Levent Akın
Bilgisayar Mühendisliği Bölümü
Boğaziçi Üniversitesi
İstanbul, Türkiye
Email: kadriye.usta,akin@boun.edu.tr

Özetçe —Robot biliminin gelişmesiyle birlikte robotlar sosyal hayatımızda yavaş yavaş yerlerini almaya başladılar. Sosyal hayatta yer alan robotlarla beraber, robot-insan ilişkisi başka bir boyuta taşınmış oldu. Bu yeni ilişkinin iyi veya kötü olacağını öngörmek günümüzde imkansızdır. İnsanlığın robotlar hakkında en büyük korkusu bir ya da bir grup robotun dünyayı ele geçirip, insanlığın sonunu getirmesidir. Bu olası felaket senaryosunu engellemek ve robot-insan ilişkisinin sınırlarını belirlemek için robot etiği bilimi ortaya çıkmıştır. Robot etiğinin nasıl olması gerektiğine dair birçok yaklaşım vardır. Biz bu çalışmamızda faydacı (utilitarianism) etik yaklaşımını izleyen kural-tabanlı bir etik birimi tasarladık. Bu birimin çıktısı olarak robot ya kontrol birimi tarafından seçilen eylemi uygulayacak ya da etik uygunluk sınırını geçtiği için eylemi iptal edip, robotu durduracaktır.

Anahtar Kelimeler—Hizmet Robotları, Robot Etiği, İnsan-Robot Etkileşimi

I. GİRİŞ

Robotların gelecekte hayatımızı etkileyeceği bir gerçek, fakat nasıl etkileyeceği konusunda hala kesin fikir üretmiyoruz; iyi ya da kötü, ikisi de olabilir. İnsanların en büyük korkusu robotların kontrolden çıkıp insanlığın sonunu getirmesidir. Peki bu durum mümkün mü? Günümüz teknolojisi ile çok zor, fakat gelecekte mümkün olabilir. Gelecek konusunda yapabileceğimiz tek şey, üretilen veya üretilecek robotların böyle bir karar vermesini engellemektir. Robot etiğinin amacı tam olarak robotların insanlığa zarar vermesini önlemektir. Robot etiği, temelini ünlü yazar Asimov'un robot yasalarına dayandırmıştır [1]. Asimov'un üç robot yasası ilk bakışta yeterli gibi görünse de sonradan kendisi de bu yasaların yeterli olmadığını anlamıştır. Bu yasaların robotlara uygulanması çok güçtür [2]. Günümüzde bir robotun hareketinin çevreye, kendine ya da başkasına etkisinin etikliğini nasıl ölçmemiz gerektiği roboetiğinin ana tartışma konusudur.

Etik birimi ile ilgili literatürde birçok bilimsel alanda (örneğin: psikoloji, sosyoloji veya hukuk) çalışma bulunmaktadır. Bunlardan en önemlisi ve robot üstünde uygulanan Winfield'in Webot'larla yaptığı, tek etik problem üzerinde yoğunlaşan çalışmadır [3]. Diğer çalışmalar ise etik birimini, biyomedikal alanında uzman sistem olarak kullanmışlardır [4], [5]. Kısacası etik birimi robot üzerinde uygulamaya konamamış ve sadece nasıl olması üzerine üretilmiş fikirler şeklinde kalmıştır. Bu alanda en kapsamlı çalışmalardan biri İngiliz Standartlar Enstitüsü'nün

Türkiye Robotbilim Konferansı, 2018

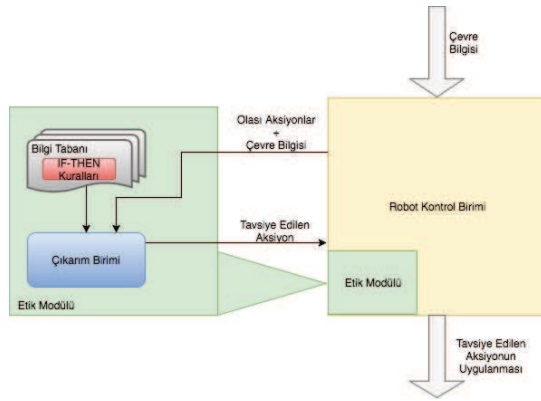
yayımladığı etik robot kılavuzudur [6]. Bu kılavuzda etik hasar tanımına, etik hasar önleme koşullarına ve etik olarak uygun birçok tasarım önerilerine yer verilmiştir.

Biz bu çalışmamızda piyasada değişik şekillerde de olsa (akıllı elektrik süpürgesi gibi) yerini almış hizmet veya servis robotları üzerinde bir etik birimi nasıl tasarlanabilir, onu inceliyoruz. Bu çalışmayla amacımız birçok etik problem karşısında tanımlanmış kurallara dayalı olarak karar veren bir birim ile robotumuzun etik davranmasını sağlamaktır. Bu etik birimi, robot kontrol biriminin üstünde uzman sistem olarak davranacak ve olası davranışların içinden etik olanı seçmeye çalışacaktır.

II. ETİK BİRİMİ

Etik davranmanın net bir tanımı bulunmamaktadır. Olaylara ve olaya bakış açısına göre bir davranışın etik olup olmadığı değişebilir. Etikçiler bile bir durumun etik olup olmadığı konusunda fikir birliğine sahip değildir. Bu durumda robot etiğine de birden fazla farklı bakış açısı vardır. Bu bakış açıları etik bilimindeki yaklaşımlardan temel alınmıştır. Etik dediğimiz kavram çok geniştir ve biyoetik, meslek etiği veya tıbbi etik gibi alt bölümlerine göre olaylara değişik yaklaşımlarda bulunur. Robot etiğinde ise iki genel yaklaşım bulunmaktadır; faydacı (utilitarianism) ve Kant'çı yaklaşım. Kant'çı yaklaşım, etik bir olayın sonucuna odaklanmaz aksine o olay sırasındaki düşünceyi önemser. Mesela boğulan bir insanı ona üzüldüğümüz için kurtarmak etik sayılmazken, yasalara göre veya evrensel ilkelere göre insan kurtarmak iyi olduğu için o insanı kurtarmak etikdir. Faydacı etik ise yapılan davranışın sonucunun insanları mutlu etme oranına göre etikliği inceler. Kısaca az önceki olayda insanın kurtarılmış olması etikdir.

Biz bu birimi kodlarken faydacı etiği kullandık çünkü çalışmamızda önemli olan robotun insana herhangi bir zarar vermeden görevini sürdürebilmesi veya yerine getirebilmesidir. Robotun görevini başarıyla bitirmesi insanı mutlu edecektir. Üstelik insana zarar vermeden (veya korkutmadan) yapması insana daha fazla fayda sağlayacaktır. Bu durumda etik birimi, robotun kontrol birimi üstünde uzman sistem olarak kullanılmalıdır. Bu birim, kontrol biriminden gelen eylemleri ve çevresel faktörleri göz önünde bulundurarak olabilecek en etik eylemi seçecek ya da hasar seviyesi belli bir seviyenin üstündeyse tamamen durmayı seçecektir. Etik biriminin genel sistemdeki yeri Şekil 1'de gösterilmiştir.



Şekil 1. Robot sistemi üzerinde etik birim şeması

A. BİRİM GEREKSİNİMLERİ

Her ne kadar robot etiğiyle ilgili birden fazla yaklaşım olsa da bu yaklaşımların hepsinin onayladığı beş tane etik birim gerekliliği bulunmaktadır. Bunlar aşağıda sıralanmıştır:

- **Tutarlılık:** Sistemde birbiri ile çelişen hiçbir kural olmamalıdır.
- **Tamlık:** Birim bir çıktı vermek zorundadır. Bu çıktı bizim tasarımımda robotun tamamen durması ya da uygun görülen davranışı yerine getirmesidir.
- **Gerçek-Zamanlılık:** Çoğu olay gibi etik problem de kısa süre içinde çözülmelidir. Eğer kullanıcının hayatı tehlikeye, hesap yapmayı bekleyecek vakti olmayacaktır.
- **Saydamlık:** Yazılan kurallar okunabilir veya dış bir uzman tarafından anlaşılabilir olmalıdır. Böylece sistemin hata ayıklaması kolaylaşacak ve seçilen eylemin etikliği kolayca sorgulanabilecektir.
- **Etik sezgiye uygunluk:** Kodlanmış kuralların belirlenen ya da belirlenecek etik kurallarına uygun olması gerekmektedir.

B. BİRİM TASARIMI

Birimimiz Şekil 1'de görüldüğü gibi elimizde olan robotun kontrol biriminin üstünde bir uzman sistem olarak iş görecektir. Etik birimimiz kural-tabanlı bir sistemdir. Bu birimin kodlanmasında CLIPS C programlama dili seçilmiştir [7]. CLIPS dilinin en büyük özellikleri taşınabilirliği ve hızıdır. Bizim çalışmamızda seçilmesinin diğer nedeni, Python dili ile bütünleşebilmesidir. Böylece gerçek dünya deneylerini gerçekleştireceğimiz NAO robotunun kontrol birimi (Python ile kodlanmış) ile iletişimde sıkıntı olmayacaktır.

Birim girdi olarak robot etrafındaki kişi sayısını, her kişi için oluşan yarar seviyesini (Beneficence), zarar seviyesini (Non-maleficence) ve kişiye sağladığı özerklik seviyesini (Autonomy) alıyor. Bu metriklerin sayısal değerleri olaylara göre oluşturulan bir tablo üzerinden alınmaktadır ve bu değerler +2 ile -2 arasında değişmektedir. Bu birimin amacı maksimum yarar ve kullanıcıya özerklik sağlarken, minimum zarar vermektir. Sonuç olarak, bu birim olabilecek en etik

davranışı kontrol birimine iletecektir. Bu değişkenlerin ayrıntılı açıklamaları aşağıda verilmiştir:

- **Özerklik Seviyesi (Autonomy):** Bu değişken robotun kullanıcının isteğine saygı duyup duymamasını inceler. Örneğin, hasta ilaç almak istemediğinde hastaya tekrar sormadan yetkili birine haber verilmesi onun özerkliği kısıtlar.
- **Yarar Seviyesi (Beneficence):** Robotun görevi başarmasının kullanıcıya sağlayacağı yararı ölçer. Eğer görev ilaç getirmekse, robotun bu ilaçları kullanıcının kişisel alanına girmeden getirmesi hem kullanıcıyı mutlu edecek hem de görev başarıyla tamamlanmış olacaktır.
- **Zarar Seviyesi (Nonmaleficence):** Tıbbi bir terimdir ve bilerek zarar vermemek anlamına gelir. Örneğin; hastaya yararlı olmayacak bir tedaviyi uygulamak. Bizim dünyamızda ise robotun kullanıcıya fazla ilaç aldırması olarak düşünülebilir.

Bu değişkenleri içeren örnek kurallarımızdan biri aşağıda gösterilmiştir:

```
if Ozerklik < 0 ve Yarar < -1 ve Zarar > 0 then
    Dur(Eylem iptal)
end if
```

III. GELECEK ÇALIŞMALAR

Etik Birimin çıkarım sistemi ve kural tabanı gerçekleştirilmiş olup belirlenen kuralları test edebilmek için NAO robot üzerinde benzetim ve gerçek hayat deneyleri yapılacaktır. Bu deney düzenekleri robotumuzun gereksinim listesinden çıkarılıp, kullanım alanında en çok karşılaşılabileceği etik ikilemlerden oluşacaktır. Bu tip etik ikileme örnek olarak şu durum gösterilebilir; hastanın alması gereken ilacı reddettiği zaman NAO'nun vereceği karar, ya hastaya ilacı alması için tekrar sormak ya da yetkili birine durumu bildirmek olabilir.

Robotumuzun seçtiği eylemin etikliği konusunda planlanan tüm deney düzenekleri ve olası robot davranışları için konunun uzmanlarına danışılacak ve sonuç tablosu oluşturulacaktır. Gerçek hayat ve benzetim deneylerinde robotumuzun davranışları gözlemlenecek, ve bu davranışlar oluşturulan tabloyla karşılaştırılacaktır.

KAYNAKÇA

- [1] G. Veruggio, J. Solis, and M. Van Der Loos, "Roboethics: Ethics applied to robotics," *IEEE Robotics and Automation Magazine*, vol. 18, no. 1, pp. 21–22, 2011.
- [2] I. Asimov, "I, robot," 2004.
- [3] A. F. T. Winfield, C. Blum, and W. Liu, "Towards an ethical robot: Internal models, consequences and ethical action selection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8717 LNAI, 2014, pp. 85–96.
- [4] M. Anderson, S. L. Anderson, and C. Armen, "An approach to computing ethics," *IEEE Intelligent Systems*, vol. 21, no. 4, pp. 56–63, 2006.
- [5] —, "MedEthEx : A Prototype Medical Ethics Advisor," *Artificial Intelligence*, no. May, pp. 1759–1765, 2006.
- [6] "Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems," British Standards Organization, Standard, 2016.
- [7] "Clips, tool for building expert systems," <http://www.clipsrules.net/>, 2013, [Online; erişim 26-Mart-2018].