

İnsansı Robotlarda Bağlamın Öğrenilmesi

Hande Çelikkanat*, Güner Orhan*, Nicolas Pugeault†, Frank Guerin‡, Erol Şahin* ve Sinan Kalkan*

*KOVAN Araştırma Laboratuvarı, Bilgisayar Mühendisliği Bölümü, Orta Doğu Teknik Üniversitesi, Ankara, Türkiye

Email: {hande,guner.orhan,erol,skalkan}@ceng.metu.edu.tr

†CVSSP, University of Surrey, UK, Email: n.pugeault@surrey.ac.uk

‡Computing Science, University of Aberdeen, Scotland, Email: f.guerin@abdn.ac.uk

Özetçe —Bu çalışmada, bağlam kavramını, bir robotun çevresiyle olan etkileşimleriyle temellendirilmiş kavramlar üzerinde tanımlamaktayız. Bu amaçla, bağlamı, berimsel dil-bilim uygulamalarında dokümanların konularını tespit etmek için yaygın olarak kullanılan Örtülü Dirichlet Paylaşımı (Latent Dirichlet Allocation) yönteminde bir örtülü değişkene karşılık gelecek şekilde modellemekteyiz. Yöntemin esnekliği, hem nesneler ve bağlamlar, hem de sahneler ve bağlamlar arasında çoktan-çoka bağıntılar tanımlanabilmesine fırsat vermektedir. Bu bağlamsal analiz, robotun algılarının üst seviye bir temsili olarak önceki çalışmalarımızda önerdiğimiz yoğun bağlantılı kavram ağı üzerinde gerçekleştirilmektedir. Ortaya çıkan bağlam bilgisi, daha sonra bu kavram ağına geri beslenerek, nesne tanıma, sahne analizi, ve eylem planlamada verimin artırılması gibi, çeşitli bilişsel fonksiyonlar dahilinde kullanılmaktadır. Sonuçlarımız, bağlamın, robotun mantık yürütmesini yönlendirmede ve bu mantık yürütme sürecinin daha verimli işlemesini sağlamakta etkili bir bilgi olduğunu göstermektedir.

Anahtar Kelimeler—Bağlam, Kavramsallaştırma, Sembol Temellendirme, Gelişimsel Robotik, Bilişsel Robotik

I. İNSANLARDA VE ROBOTLARDA BAĞLAM

Bağlam kavramı, bir duruma dair, onu diğer durumlardan ayıran bilgilerin tamamını belirtir. İnsanların hemen hemen her konuda, duyumsal bilgileri tek başlarına ayrı ayrı incelemek yerine, bir bağlam çerçevesi içerisinde yorumladıkları birbirlerinden bağımsız senaryolarda gösterilmiştir [1]. Bağlamın etkisi özellikle gördüklerimizi yorumlamada [2], [3] ve dili kullanmada [4] belirgin olsa da, hiçbir bilişsel fonksiyon tamamıyla bağlamdan bağımsız sayılmaz[5]. Örneğin insanların günlük hayatta sıkça karşılaştıkları durumlara hızlı tepki verebilmek için önceden çıkardıkları betikleri kullandıkları bilinmektedir [6]. Bu çalışmada, bir robotun, bağlamı, gördüklerini doğru yorumlamak ve etkin planlama yapmak için kullanabileceği gösterilmektedir.

II. YÖNTEM

Bu çalışma iCub insansı robot platformu üzerinde gerçekleştirilmiştir. Robot önündeki masanın üzerinde bulunan nesnelerle, dokunarak, sallayarak, vb. etkileşimde bulunabilmektedir. Bu etkileşim sonucunda her nesneden 92 tane fiziksel özellik elde edilmektedir. Bunlardan 67 tanesi Kinect 3B kamerası ile algılanan yüzey normalleri gibi görsel özellikler, 13 tanesi nesneye ait ses verileri, 6 tanesi parmaklardaki dokunma sensörleriyle algılanan sertlik bilgileri ve 6 tanesi de robot kolunun eklemlerinin durum bilgileridir. Nesne seti 15'i test nesnesi olmak üzere 60 nesneden oluşmaktadır. Bu nesneler 6 isim kümesine ($\mathcal{I} =$

Tablo I. Bağlam terimlerinin denklikleri

Örtülü Dirichlet Paylaşımı	Bu Çalışma
doküman $d \in \mathcal{D}$	tek bir sahne (kavram ağı olarak ifade edilerek)
korpus $\mathcal{D}$	öğrenme fazı sırasında karşılaşılan bütün sahneler
kelime $w_i \in \mathcal{W}$	kavram ağı içindeki aktif bir kavram, c_{act}
	(isim, sıfat, ya da eylem olabilir: $c_{act} \in \mathcal{I} \cup \mathcal{S} \cup \mathcal{E}$)
konu $t_i \in \mathcal{T}$	bir bağlam, Mutfak, Oyun odası, ya da Atölye olabilir

{top, silindir, kutu, bardak, tabak, alet} ve 10 sıfat kümesine ($\mathcal{S} = \{köşeli, yuvarlak, uzun, kısa, ince, kalın, sesli, sessiz\}$) bölünmektedir. Robot nesneler üzerinde 7 çeşit eylem uygulayabilmektedir: $\mathcal{E} = \{taşı, it, tut, devir, salla, bırak, at\}$. Kavramlar prototipleriyle ifade edilmektedirler [7]. İlk defa karşılaşılan bir nesne, özellik vektörü kavramların prototipleriyle karşılaştırılarak uygun kavramlarla ilişkilendirilir.

Kavram Ağı Robot karşılaştığı nesnelere ait kavramları bir kavram ağı içinde değerlendirmektedir [8]. Bu ağ, düğümleri kavramlara, düğümler arası bağlantıları ise kavramlar arası ilişkilere denk gelen bir Rastgele Markov Alanı (ing. Markov Random Field) olarak modellenmiştir. Markov alanının bağlantıları, öğrenme kümesi içindeki nesnelerde bir arada bulunan kavramlar birbirine bağlı olacak şekilde ayarlanır. Daha sonra yeni bir nesneyle karşılaşıldığında, prototipler üzerinden belirlenen ilgili kavramlar, kavram ağına taşınır, ve Döngüsel Fikir Aktarımı (ing. Loopy Belief Propagation) yöntemiyle aktivasyonlarını bu ağ üzerinde yayarlar. Kavramlar arası bağlantılar sayesinde, bu aktivasyonlar, önceden öğrenilmiş olan, sıkça birlikte aktive olan kavramlar bilgisine en uygun şekilde yayılırlar. Neticede elde edilen kavram ağının, minimum enerjisine denk gelen son durumu, hem anlık algılara, hem de önceden edinilmiş bilgiye en uygun tahminleri verir.

Örtülü Dirichlet Paylaşımı ile Bağlam Çıkarımı Bağlamsal çıkarım, kavram ağı üzerinden Örtülü Dirichlet Paylaşımı (ing. Latent Dirichlet Allocation) [9] yöntemi ile gerçekleştirilir. Bu yöntem esasen metin dokümanlarının konularını belirlemek için önerilmiş üretici (ing. generative) ve güdümsüz (ing. unsupervised) bir yöntemdir. Dokümanlar belli bir konu setindeki konulara ait olmalarının olasılık dağılımıyla, konular ise belli bir kelime setindeki kelimeleri üretmelerine dair olasılık dağılımıyla tanımlanır. Terimlerin sistemimizdeki denklikleri Tablo I'de gösterilmiştir.

Bağlamın Kavram Ağı İçerisinde Kullanılması Kavram ağı üzerinde Örtülü Dirichlet Paylaşımı yöntemi ile çıkarılan bağlam, kavram ağına geri beslenerek ve Rastgele Markov Alanının minimum enerji durumunu yeniden hesaplanarak tahminlerin güncellenmesinde kullanılır.

Tablo II. Nesnelerin tanınması. Sadece prototipler kullanılarak, kavram ağı kullanılarak, ve bağlam kullanılarak elde edilen sonuçlar karşılaştırılmıştır. Doğru tahminler kalın fontla, yanlış tahminler üstü çizilerek belirtilmiştir.

Sadece Algı			Kavram Ağı			Bağlam				
Nesneler	İsimler (% kesinlik)	Sıfatlar (% kesinlik)	İsimler (% kesinlik)	Sıfatlar (% kesinlik)		Sahne	İsimler (% kesinlik)	Sıfatlar (% kesinlik)		
	top (8%) kutu (13%) bardak (43%) silindir (20%) tabak (9%) alet (7%)	köşeli (34%) sert (71%) sesli (42%) kısa (54%) uzun (46%) kalın (47%)	yuvarlak (66%) yumuşak (29%) sessiz(58%) ince (53%)	top (0%) kutu (0%) bardak (100%) silindir (0%) tabak (0%) alet (0%)	köşeli (0%) sert (100%) sesli (0%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(100%) ince (0 %)		top (0%) kutu (0%) bardak (100%) silindir (0%) tabak (0%) alet (0%)	köşeli (0%) sert (100%) sesli (0%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(100%) ince (0 %)
	top (33%) kutu (16%) bardak (13%) silindir (13%) tabak (14%) alet (11%)	köşeli (42%) sert (39%) sesli (62%) kısa (61%) uzun (39%) kalın (56%)	yuvarlak (58%) yumuşak (61%) sessiz(38%) ince (44 %)	top (100%) kutu (0%) bardak (0%) silindir (0%) tabak (0%) alet (0%)	köşeli (0%) sert (0%) sesli (100%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (100%) yumuşak (100%) sessiz(0%) ince (0 %)		top (100%) kutu (0%) bardak (0%) silindir (0%) tabak (0%) alet (0%)	köşeli (100%) sert (0%) sesli (0%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (0%) yumuşak (100%) sessiz(100%) uzun (0%) ince (0 %)
	top (12%) kutu (13%) bardak (17%) silindir (29%) tabak (12%) alet (17%)	köşeli (45%) sert (56%) sesli (58%) kısa (41%) uzun (59%) kalın (40%)	yuvarlak (55%) yumuşak (44%) sessiz(42%) uzun (59%) ince (60%)	top (0%) kutu (0%) bardak (0%) silindir (100%) tabak (0%) alet (0%)	köşeli (0%) sert (0%) sesli (100%) kısa (100%) uzun (100%) kalın (0%)	yuvarlak (100%) yumuşak (0%) sessiz(0%) ince (100%)		top (0%) kutu (0%) bardak (0%) silindir (100%) tabak (0%) alet (0%)	köşeli (0%) sert (100%) sesli (100%) kısa (0%) uzun (100%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(0%) uzun (100%) ince (0 %)
	top (14%) kutu (43%) bardak (12%) silindir (12%) tabak (11%) alet (8%)	köşeli (64%) sert (34%) sesli (30%) kısa (59%) uzun (41%) kalın (63%)	yuvarlak (36%) yumuşak (66%) sessiz(70%) uzun (41%) ince (37 %)	top (0%) kutu (100%) bardak (0%) silindir (0%) tabak (0%) alet (0%)	köşeli (100%) sert (0%) sesli (0%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (0%) yumuşak (100%) sessiz(100%) uzun (0%) ince (0 %)		top (0%) kutu (100%) bardak (0%) silindir (0%) tabak (0%) alet (0%)	köşeli (100%) sert (0%) sesli (0%) kısa (100%) uzun (0%) kalın (100%)	yuvarlak (0%) yumuşak (100%) sessiz(100%) uzun (0%) ince (0 %)
	top (12%) kutu (13%) bardak (12%) silindir (14%) tabak (40%) alet (9%)	köşeli (46%) sert (53%) sesli (44%) kısa (45%) uzun (55%) kalın (59%)	yuvarlak (54%) yumuşak (47%) sessiz(56%) uzun (55%) ince (41%)	top (0%) kutu (0%) bardak (0%) silindir (0%) tabak (100%) alet (0%)	köşeli (0%) sert (100%) sesli (100%) kısa (0%) uzun (100%) kalın (100%)	yuvarlak (0%) yumuşak (0%) sessiz(100%) uzun (100%) ince (0 %)		top (0%) kutu (0%) bardak (0%) silindir (0%) tabak (100%) alet (0%)	köşeli (0%) sert (100%) sesli (0%) kısa (0%) uzun (100%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(100%) uzun (100%) ince (0 %)
	top (11%) kutu (13%) bardak (15%) silindir (18%) tabak (11%) alet (32%)	köşeli (48%) sert (55%) sesli (61%) kısa (39%) uzun (61%) kalın (57%)	yuvarlak (52%) yumuşak (45%) sessiz(39%) uzun (61%) ince (43 %)	top (0%) kutu (0%) bardak (0%) silindir (0%) tabak (0%) alet (100%)	köşeli (0%) sert (100%) sesli (100%) kısa (0%) uzun (100%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(0%) uzun (100%) ince (0 %)		top (0%) kutu (0%) bardak (0%) silindir (0%) tabak (0%) alet (100%)	köşeli (0%) sert (100%) sesli (100%) kısa (0%) uzun (100%) kalın (100%)	yuvarlak (100%) yumuşak (0%) sessiz(0%) uzun (100%) ince (0 %)
	top (14%) kutu (17%) bardak (19%) silindir (26%) tabak (13%) alet (11%)	köşeli (42%) sert (60%) sesli (42%) kısa (45%) uzun (55%) kalın (40%)	yuvarlak (58%) yumuşak (40%) sessiz(58%) uzun (55%) ince (60 %)	top (0%) kutu (0%) bardak (100%) silindir (0%) tabak (0%) alet (0%)	köşeli (0%) sert (100%) sesli (0%) kısa (100%) kalın (0%)	yuvarlak (100%) yumuşak (0%) sessiz(100%) uzun (0%) ince (100 %)		top (0%) kutu (0%) bardak (0%) silindir (100%) tabak (0%) alet (0%)	köşeli (0%) sert (100%) sesli (0%) kısa (100%) kalın (0%)	yuvarlak (100%) yumuşak (0%) sessiz(100%) uzun (0%) ince (100 %)

Tablo III. Sahnelerin tanınması. Doğru tahminler kalın fontla, yanlış tahminler üstü çizilerek belirtilmiştir.

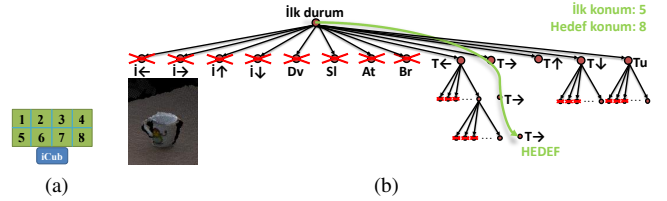
Tekil Bağlamlar			Çoklu Bağlamlar		
Nesneler	Tahmin Edilen Bağlam (% kesinlik)		Sahne	Nesneler	Tahmin Edilen Bağlam (% kesinlik)
	2 bardak, 2 tabak	Mutfak (100%)		3 kutu, 1 top, 1 silindir, 1 alet	Oyun odası (72.59%) Atölye (26.23%) Mutfak (1.18%)
	2 kutu, 2 top	Oyun odası (100%)		2 tabak, 2 bardak, 1 kutu 1 top,	Mutfak (62.04%) Oyun odası (37.14%) Atölye (0.82%)
	3 silindir 2 alet	Atölye (100%)		1 alet, 1 silindir, 1 tabak, 1 bardak	Mutfak (46.67%) Atölye (51.56%) Oyun odası (1.77%)

III. SONUÇLAR

Tablo II ve III, yöntemin nesne tanıma ve sahne analizindeki performansını göstermektedir. Tablo II'de çeşitli nesnelerin tanınması, sadece prototiplerin kullanılması durumunda, kavram ağının kullanılması durumunda, ve bağlam bilgisinin de hesaba katılması durumunda karşılaştırılmıştır. Sonuçlar bağlam bilgisinin yanlış tahminleri düzeltmekte etkili olduğunu göstermektedir. Tablo III sistemin tekli ve çoklu bağlamlardan oluşan sahneleri başarılı bir biçimde ayırdı edip yorumlayabildiğini göstermektedir. Figür 1'de ise bağlam bilgisi (Mutfak) bir bardağın yerinin değiştirilmesine ilişkin planlama ağacının budanmasında (tehlikeli davranışları, *örn.* itme, budayarak) kullanılmıştır.

TEŞEKKÜR

Bu çalışma, 111E287 nolu TÜBİTAK projesi kapsamında olarak desteklenmiştir.



Şekil 1. Planlama ağacının bağlam bilgisiyle budanması. İ←: Sola it, İ→: Sağa it, İ↑: İleri it, İ↓: Geri it, Dv: Devir, Sl: Salla, At: At, Br: Bırak, T←: Sola taşı, T→: Sağa taşı, T↑: İleri taşı, T↓: Geri taşı, Tu: Tut.

KAYNAKÇA

- [1] J. A. Bargh, M. Chen, and L. Burrows, "Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action." *Journal of personality and social psychology*, 1996.
- [2] M. Bar, "Visual objects in context," *Nature Reviews Neuroscience*, vol. 5, no. 8, pp. 617–629, 2004.
- [3] A. Torralba, "Contextual priming for object detection," *Int. Journal of Computer Vision*, vol. 53, no. 2, pp. 169–191, 2003.
- [4] T. A. Van Dijk, "Discourse and context," *A Sociocognitive Approach*. Cambridge, 2008.
- [5] W. Yeh and L. W. Barsalou, "The situated nature of concepts," *The American journal of psychology*, pp. 349–384, 2006.
- [6] R. C. Schank and R. P. Abelson, "Scripts, plans, goals and understanding," *Hillsdale, NJ: Lawrence Erlbaum*, 1977.
- [7] S. Kalkan, N. Dag, O. Yürten, A. M. Borghi, and E. Sahin, "Verb concepts from affordances," *Interaction Studies*, vol. 15, 2014.
- [8] G. Orhan, H. Çelikkanat, and S. Kalkan, "A probabilistic web of concepts on a humanoid robot," submitted.
- [9] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, 2003.