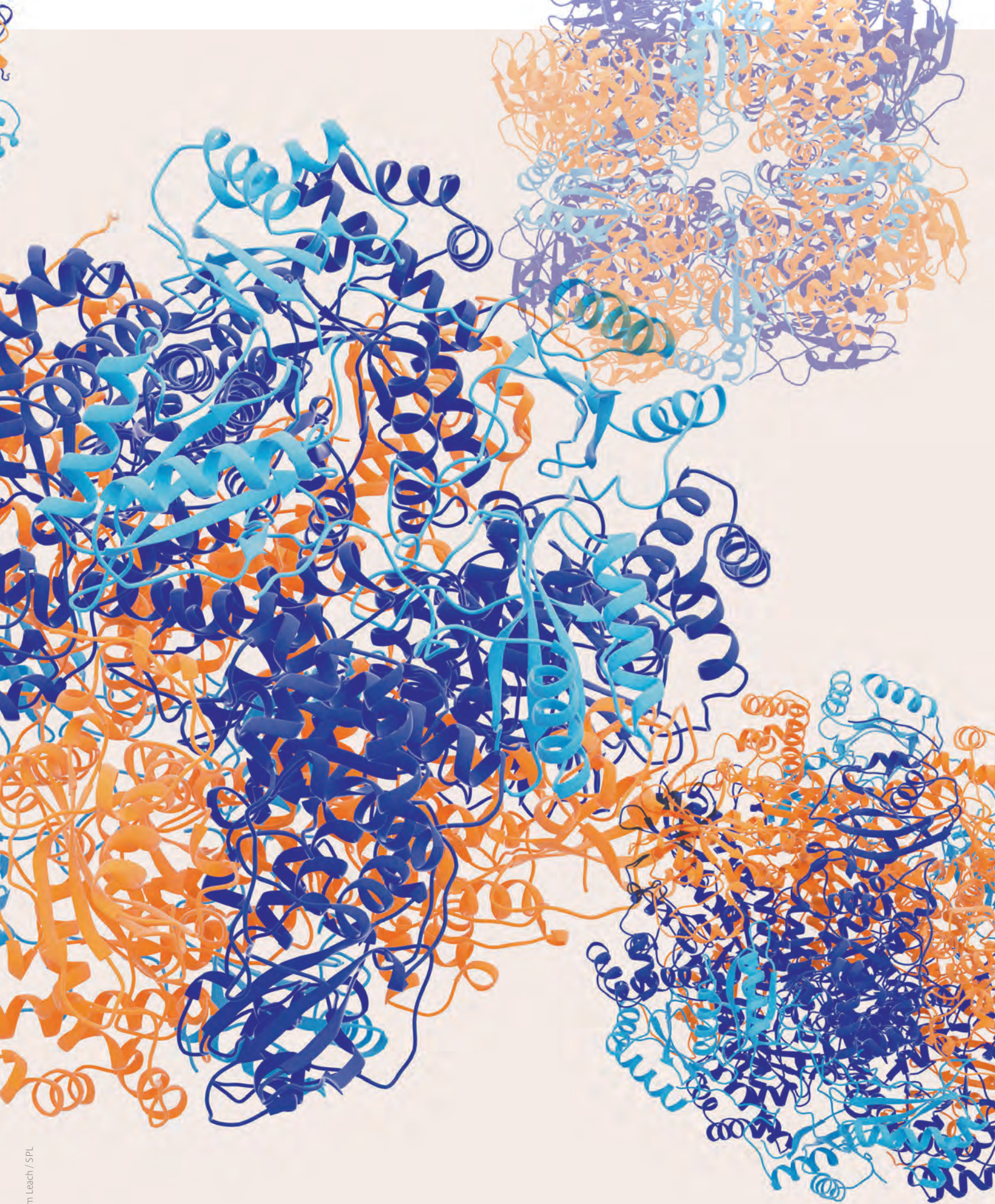


Biyolojik Moleküller, Topoloji ve Yapay Zekâ

Dr. Mahir E. Ocak [TÜBİTAK Bilim ve Teknik Dergisi

Proteinler ve DNA gibi biyolojik moleküllerdeki düğümler üzerine çalışmalar yapan araştırmacılar, yüz yıldan uzun bir süredir çözilemeyen bir matematik problemiyle karşı karşıya kalıyor. Son bilimsel çalışmalar, yapay zekânın bu sorunun çözümünde de yararlı olabileceğini gösteriyor.





Düğümlü Proteinler

Proteinler, binlerce atomdan oluşan makromoleküllerdir. Amino asitlerin art arda eklenmesiyle oluşan bu biyolojik moleküller önce düz bir zincir biçiminde sentezlenir daha sonra katlanarak karmaşık geometrik şekillere girer. Sahip oldukları bu karmaşık yapılar, proteinlerin vücutta

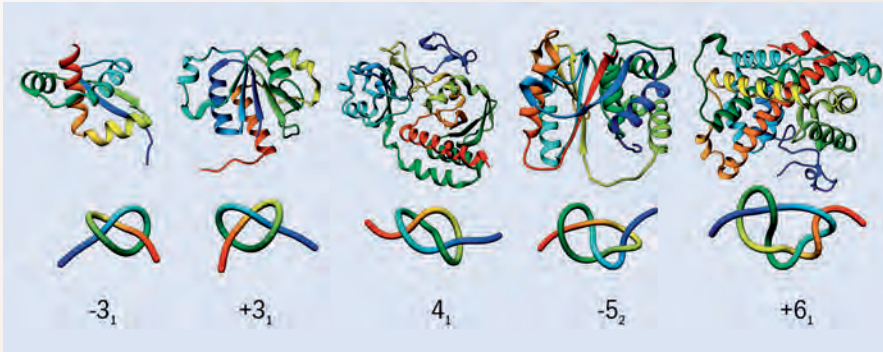
kimyasal tepkimeleri hızlandırma, molekülleri bir yerden başka bir yere aktarma ya da uyarılara tepki verme gibi işlevleri yerine getirmelerini sağlar.

Her ne kadar proteinler katlanıp şekil alırken amino asit zincirinin farklı bölgeleri birbirine dolansa da düğümler oluşması beklenen bir durum değildir. Ancak 1970'lerde yapılan bilimsel çalışmalar, pek çok proteinin

yapısında düğümler olduğunu gösterdi. Bugün Protein Veri Bankası'nda yer alan proteinlerin yaklaşık %1'inin düğümlü yapıda olduğu biliniyor.

Düğümlerin varlığı proteinlerin katlanma sürecine, kararlılığına ve işleyişine olumlu ya da olumsuz etki edebilir. Örneğin, düğümlü proteinlerin düğümsüz proteinlere kıyasla aşırı sıcaklıklara daha dayanıklı olduğunu gösteren çalışmalar var. Ancak düğümlü proteinlerin katlanma süreci muhtemelen daha yavaş gerçekleşiyor ve bu süreçte hataların ortaya çıkma ihtimali daha yüksek oluyor.

Proteinlerin yapılarındaki düğümlerin nasıl ve neden oluştuğu bugün tam olarak bilinmiyor. Ayrıca düğümlerin, proteinlerin biyolojik işlevleri üzerindeki etkilerinin ne olduğu da hâlâ açıklanmayı bekliyor. Bu ve benzeri sorulara cevap bulabilmek için yapılması gerekenlerin başında proteinlerin yapılarındaki düğümleri sınıflandırmak geliyor. Bu durum da araştırmacıları yüz yıldan uzun bir süredir cevap bulunamayan bir matematik problemiyle karşı karşıya bırakıyor: Farklı düğümlerin birbirinden ayırt edilmesini sağlayacak bir topolojik değişmezin keşfedilmesi.



Çeşitli düğümlü protein yapıları. Görsellerin altındaki sayılar, yapıdaki kesişim noktalarının sayısını ifade ediyor. Sayıların başındaki + ve - işaretleri, yapıların birbirinin ayna görüntüsü olduğu anlamına geliyor.

Topolojik Değişmezler

Biyolojik moleküller için “bir topolojik değişmez” keşfetme çabalarına göz atmadan önce topolojik değişmez kavramına odaklanalım.

Matematiğin topoloji olarak adlandırılan dalı, geometrik nesnelerin ufak deformasyonlar altında değişmeyen özellikleriyle ilgilenir. İki geometrik cismin topolojik olarak birbirine denk olması kesme, koparma, yapıştırma gibi işlemlere başvurmaksızın sadece esnetme, sıkıştırma gibi ufak deformasyonlarla bir cismin diğerine dönüştürülebileceği anlamına gelir. Topolojik değişmezler ise topolojik olarak farklı geometrik nesneleri ayırt etmeye yarar. İki geometrik cisim için bir topolojik değişmezin değerinin aynı olması, cisimlerin topolojik olarak birbirine denk olduğu yani ufak deformasyonlarla birinin diğerine dönüştürülebileceği anlamına gelir. Topolojik değişmez değerinin birbirinden farklı olduğu iki geometrik cisim ise ufak deformasyonlarla birbirine dönüştürülemez.

Topolojik değişmez kavramını iki boyutlu yüzeyler üzerinden örneklendirelim. İki boyutlu yüzeyler, “genus” olarak adlandırılan bir topolojik değişmeze sınıflandırılır. Genus, yüzeydeki



Robert Brook / SPL

Genusu 1 olan çeşitli iki boyutlu yüzeyler. Genus, yüzeydeki kulpların sayısıdır. Aynı genusa sahip olmaları, tüm bu yapıların ufak deformasyonlarla birbirine dönüştürülebilecekleri anlamına gelir.

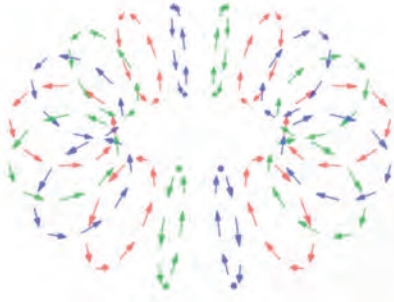
deliklerin (ya da kulpların) sayısıdır. Örneğin, bir kürenin genusu sıfırdır. Çünkü kürenin yüzeyinde hiç delik yoktur. Bir torusun genusu ise 1’dir. Çünkü torusun ortasında 1 tane delik vardır. Genusların farklı olması, kürenin ve torusun topolojik olarak farklı yapıda oldukları anlamına gelir. Bir torus, ufak deformasyonlarla bir küreye dönüştürülemez. Bir torusu farklı yönlerde ne kadar esnetirseniz esnetin ne kadar sıkıştırırsanız sıkıştırın her zaman bir tane deliği vardır.

Yüzeydeki kulp sayısının bir topolojik değişmez olması, aynı sayıda kulpa sahip geometrik şekillerin topolojik olarak birbirine denk oldukları anlamına gelir. Örneğin, her ikisinin de sadece bir kulpu olduğu için bir torus ve bir kupanın yüzeyi, topolojik olarak birbirine denktir. Bir torus, ufak deformasyonlarla bir kupanın yüzeyine dönüştürülebilir.

Düğüm Teorisi

Modern düğüm teorisinin 1867 yılında Peter Guthrie Tait’in William Thomson’a (Lord Kelvin) yaptığı bir gösteri deneyiyle başladığı söylenir. Tait, deneyler sırasında duman halkaları üretmişti. İlginç olan ise duman halkalarının yapısının şaşırtıcı derecede kararlı olmasıydı. Duman halkaları, odanın içindeki havada yapıları bozulmadan yol alıyor ve hatta sanki elastik bir malzemeden yapılmış gibi birbirleriyle çarpışarak geri yansıyor.

Duman halkaları, vorteks halkalarının örnekleridir. Bu halkalarda aerosoller ve partiküller, ufak çembersel rotalar takip eder (bkz. 34. sayfa, sol üstteki grafik). Duman halkalarının kararlılığını sağlayan, aerosollerin ve partiküllerin bu hareketidir.



Duman halkaları, vorteks halkalarının bir örneğidir.

Kelvin, tanık olduğu deneyden esinlenerek bir atom teorisi öne sürdü. O yıllarda uzayın "eter" adı

verilen bir madde ile dolu olduğu, elektromanyetik dalgaların bu maddenin içinde yol aldığı düşüncesi hakimdi. Kelvin, atomların eter içindeki halkalar olduğunu ve yapılarının duman halkalarınınkine benzediğini iddia etti. Farklı tür atomları birbirinden ayırt eden şeyin ise bu vorteks halkalarındaki düğümler olduğunu öne sürdü.

Kelvin'in atom teorisi, Tait'in ilgisini çekmişti. Tait, farklı tür düğümleri kataloglamaya başladı. Takip eden 25 sene içinde en fazla on kesişim noktasına sahip düğümler içeren bir tablo oluşturdu. Tait'in tablosunda 251 ayrı çizim vardı. İlerleyen yıllarda Kelvin'in atom teorisi çürütüldü ancak düğüm teorisi, matematiğin bir dalı olarak gelişmeye devam etti.

Düğümler ve Topolojik Değişmezler

Bugüne kadar düğümleri sınıflandırmak amacıyla pek çok topolojik değişmez adayı öne sürüldü. Ancak hiçbirisi tam anlamıyla bir topolojik değişmez değil. Görece basit yapıdaki düğümleri sınıflandırmak için yararlı olsalar da her durumda doğru sonuç vermiyorlar.

Bugüne kadar öne sürülen topolojik değişmezleri cebirsel ve geometrik olarak iki gruba ayırabiliriz. Cebirsel değişmezler düğümlerin "iki boyutlu izdüşümleri" üzerinden; geometrik değişmezlerse düğümlerin üç boyutlu yapısı üzerinden hesaplanır.

Geçmişte öne sürülmüş cebirsel değişmezlerden biri Dowker-Thistlethwaite kodu. 1983 yılında öne sürülmüş bu değişmez şöyle hesaplanıyor: 1) n tane kesişim noktası olan bir düğüm boyunca yol olarak her bir kesişim noktasını 1, 2, 3 ... $2n$ sayılarıyla etiketleyin. Sonuçta

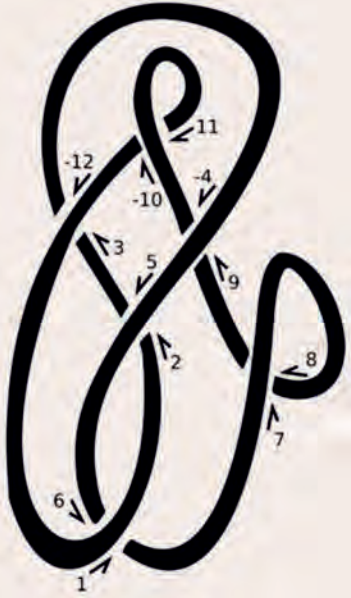
THE FIRST SEVEN ORDERS OF KNOTTINESS.

Plate VI



Tait'in 1884'te yayımladığı "On Knots Part II" adlı çalışmasında yer alan, yediye kadar kesişim noktasına sahip düğümlerin kataloglandığı bir tablo.

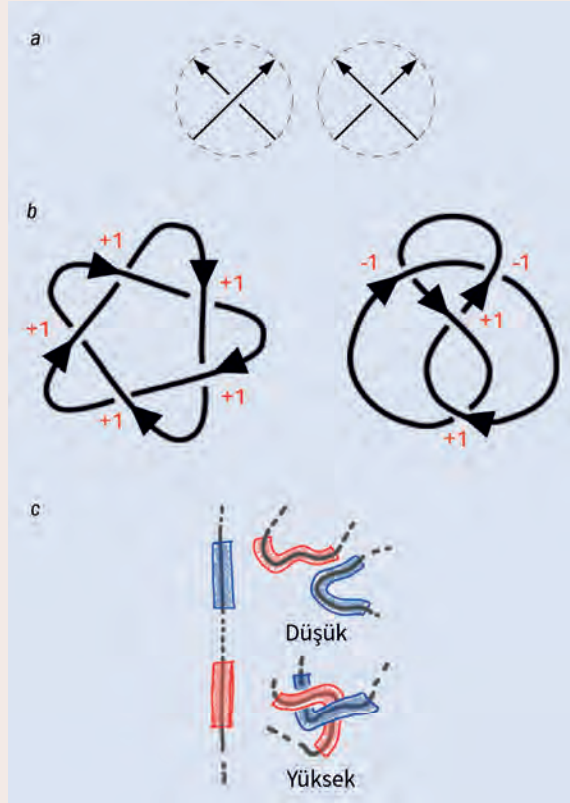
her bir kesişim noktası, biri tek, diğeri çift olmak üzere bir sayı çifti ile etiketlenenecektir. 2) Eğer sayı çifti ve etiket bir üst kesişim noktasına aitse sayının işaretini negatif yapın. Elde edeceğiniz çift sayı dizisi, düğümü tanımlayan bir kod olacaktır.



Dowker-Thistlethwaite kodunun elde edilmesine dair bir örnek.

Tait'in kendisi de kıvrılma sayısı olarak adlandırdığı bir değişmez öne sürmüştü. Kıvrılma sayısını hesaplamak için düğümdeki her bir kesişim noktası, ilk olarak üst kesişim noktası ya da alt kesişim noktası olarak sınıflandırılıyor (bkz. sağ üstteki şekil). Daha sonra alt kesişim noktaları -1, üst kesişim noktaları +1 ile etiketleniyor. Son olarak bu sayılar toplanıyor.

Kıvrılma sayısını sadece cebirsel olarak değil, proteinler gibi moleküllerin üç boyutlu yapıları üzerinden geometrik olarak hesaplamak da mümkün. Ayrıca



(a) Kıvrılma sayısı hesaplanırken her bir kesişim noktası, üst kesişim noktası (solda) ya da alt kesişim noktası (sağda) olarak sınıflandırılır. (b) Üst kesişim noktaları +1 ile alt kesişim noktaları -1 ile etiketlenir. (c) Kıvrılma sayısı, geometrik olarak da hesaplanabilir.

geometrik kıvrılma sayısı, tüm moleküler yapının yanı sıra molekülün belirli bir bölgesi için de hesaplanabilir. “Yerel” kıvrılma sayısı, molekülün belirli bölgelerinin birbirine ne ölçüde dolandığını gösterir.

Moleküler Düğümler ve Yapay Zekâ

Yapay zekâ gün geçtikçe bilimsel araştırmalarda daha da fazla yer alıyor. Söz konusu biyolojik moleküller olduğunda da yapay zekâ önemli bir yardımcı hâline geldi. Örneğin Google DeepMind’ın AlphaFold yapay zekâ programı

2021 yılında, çok uzun bir süredir üzerine yoğun araştırmalar yapılan bir konuda önemli bir başarıya imza attı. AlphaFold, amino asit dizilimlerinden proteinlerin yapısını büyük bir doğrulukla tahmin etmeyi başarıyor. Bir proteinin işlevi, üç boyutlu yapısına bağlıdır. Bu durum AlphaFold’un özellikle yeni ilaç geliştirme çabalarında önemli bir araç hâline gelmesine yol açtı. AlphaFold’u geliştiren araştırmacılar 2024 Nobel Kimya Ödülü ile onurlandırıldı.

Son yıllarda bazı araştırmacılar, yapay zekâyı düğüm değişmezi probleminde de kullanmaya başladı. Hong Kong Şehir Üniversitesinden Olafs Vandans ve arkadaşları 2020’de, Padova Üniversitesinden Anna Braghetto ve arkadaşları 2023’te yayımladıkları makalelerde yapay zekâ kullanarak düğümleri sınıflandırma üzerine çalışmalar yaptı. Her iki araştırma grubu da düğümleri art arda dizilmiş boncuklardan oluşan sicimler gibi modelledi ve kartezyen koordinatlar kullanarak yapay zekâ programını eğitti. Padova ve arkadaşları, farklı olarak eğitimler sırasında yapay zekâ uygulamasına boncuklar arasındaki mesafe ve açı ile ilgili bilgiler de verdi. Her iki araştırma grubu da büyük bir doğrulukla düğümleri

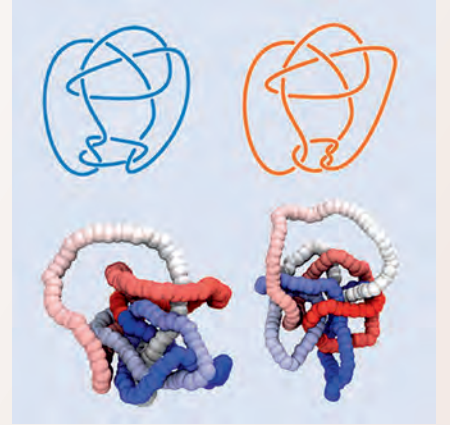
sınıflandırmayı başardı. Ancak çalışmalar, en basit yapıdaki beş ayrı düğümle sınırlıydı.

Yakın zamanlarda Edinburgh Üniversitesinden bir grup araştırmacı, yapay zekâ kullanarak en fazla on kesişim noktasına sahip düğümleri %95 doğrulukla sınıflandırmayı başardı. Çalışmaya imza atan Davide Michieletto ve öğrencileri, düğümlü proteinlerle ilgili biyolojik süreçlerden esinlendi.

DNA moleküllerindeki düğümleri çözmek için özelleşmiş “topoizomeraz” olarak adlandırılan enzimler vardır. Bu enzimler, DNA iplikçiklerini kesip yeniden yapııştırarak düğümleri çözer. Bu süreçlerle ilgili dikkat çekici bir nokta ise topoizomerazların DNA moleküllerinden binlerce kat daha küçük olmasıdır. Bu durumdan yola çıkan

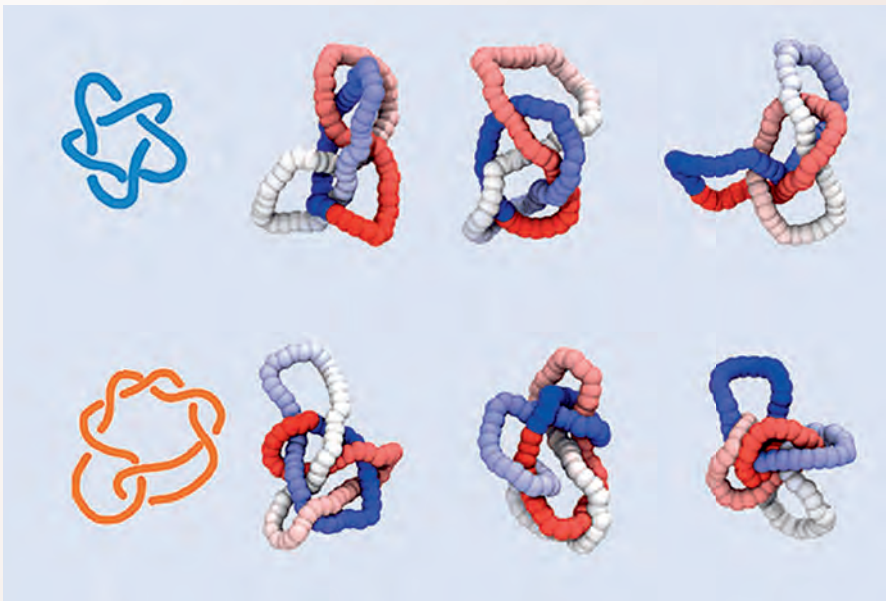
araştırmacılar, topoizomerazların DNA moleküllerinin yerel geometrik özelliklerine duyarlı olduğunu varsaydı.

Araştırmacılar, iki ayrı yapay zekâ uygulaması geliştirdi. Uygulamaların birini daha önceki çalışmalarda olduğu gibi kartezyen koordinatlarla, diğerini ise yerel geometrik kıvrılma sayısı ile eğitti. Eğitimler sırasında en basit beş düğümün termal hareketlerinin simüle edilmesiyle elde edilmiş binlerce farklı konfigürasyon içeren bir veri seti kullanıldı. Testler sırasında ise yapay zekâ uygulamasından veri tabanında yer almayan düğüm konfigürasyonlarını sınıflandırması istendi. Kartezyen koordinatlarla eğitilen yapay zekâ uygulaması, her beş konfigürasyondan dördünü doğru sınıflandırabilirdi. Yerel kıvrılma sayısı ile eğitilen yapay zekâ uygulamasının doğruluk oranı ise %99,9 oldu.



Conway ve Kinoshita-Terasaka düğümleri. Pek çok topolojik değişmez, yapıları birbirine çok benzeyen bu iki düğümü ayırt edemiyor.

Araştırmacılar, bir başka testte ise uygulamalardan üç ayrı düğümü sınıflandırmasını istedi. Bu düğümlerden Conway düğümü ve Kinoshita-Terasaka düğümü olarak adlandırılan ikisi, on bir kesişim noktasına sahipti. Testte kullanılan üçüncü düğüm ise tüm düğümlerin en basiti olan ve hiçbir düğüm içermeyen basit bir halkaydı. Conway düğümü ve Kinoshita-Terasaka düğümünün yapıları birbirine çok benziyor ve bugüne kadar öne sürülmüş topolojik değişmezlerin pek çoğu, bu iki düğümü birbirinden ayırt etmeyi başaramıyor. Bu düğümlerin topolojik değişmezlerinin bazıları da basit halkayla aynı değerlere sahip. Elde edilen sonuçlar, kartezyen koordinatlarla eğitilen yapay zekâ uygulamasının Conway düğümü ile Kinoshita-Terasaka düğümünü birbirinden ayırt etmeyi başaramadığını gösterdi. Yerel geometrik kıvrılma



Yapay zekâyı eğitmek için kullanılan düğümlerden ikisi ve bu düğümlerin farklı konfigürasyonları

sayısı ile eğitilen yapay zekâ uygulaması ise kendisine verilen konfigürasyonların %99,9'unu doğru biçimde sınıflandırmayı başardı.

Araştırmacılar, gerçekleştirdikleri son testte ise çok daha büyük bir veri tabanı kullandı. İlk olarak her biri en fazla 10 kesişim noktasına sahip 250 ayrı düğümün termal hareketleri simüle edildi. Daha sonra yapay zekâ uygulamalarından kendilerine verilen düğümleri sınıflandırmaları istendi. Kartezyen koordinatlarla eğitilen yapay zekâ uygulaması, bu karmaşık yapıları düğümleri sınıflandırmakta pek başarılı olmadı. Uygulama, ortalama olarak her beş testten sadece birinde doğru sonuç verebildi. Yerel geometrik kıvrılma sayısı ile eğitilen yapay zekâ uygulamasının doğruluk oranı ise %95 oldu. Uygulamanın ulaştığı bu doğruluk oranına hiçbir topolojik değişmez tek başına ulaşamıyor.

Araştırmacıların geliştirdikleri yapay zekâ uygulaması üzerinde yaptıkları analizler, yapay zekânın bu kadar büyük bir doğruluğa ulaşmayı nasıl başardığını gösteriyor. Uygulama, beş basit

düğüm ile yapılan testler sırasında düğüm üzerindeki her bir nokta çiftini alıp iki noktadaki kıvrılma değerlerini birbiriyle çarpıyor. İlginç bir nokta ise bu prosedürün “ikinci dereceden Vassiliev değişmezi” olarak adlandırılan bir değişmezi hesaplarken kullanılması. Vassiliev değişmezlerinin farklı dereceleri, yerel kıvrılma matrisi çiftlerini, üçlülerini, dörtlülerini (...) çarparak hesaplanıyor. Geçmişte tüm Vassiliev değişmezleri kümesinin, düğümleri sınıflandırmak için kullanılabilecek gerçek bir topolojik değişmez olduğu öne sürülmüş ancak bu önermeyi ispatlamak mümkün olmamıştı. Daha da ilginç olanı, topolojisi giderek karmaşılaşan düğümler verildiğinde yapay zekâ uygulamasının giderek daha yüksek dereceden Vassiliev değişmezlerini hesaplamaya başlaması. Örneğin 250 düğümle ilgili bilgileri içeren veri tabanı kullanıldığında uygulama, üçüncü ya da dördüncü dereceye kadar Vassiliev değişmezlerini hesaplamaya başlıyor. Ne düğümler ne de düğüm teorisi hakkında hiçbir bilgisi olmayan bir programın kendini zorlu problemlere adapte edebilmesi, yapay zekânın potansiyelini gözler önüne seriyor.

Özet

Farklı düğümleri sınıflandırmak için kullanılacak bir topolojik değişmez keşfetmek, yüz yıldan uzun bir süredir çözülemeyen bir matematik problemi. Son zamanlarda proteinler ve diğer biyolojik moleküllerdeki düğümler üzerine çalışmalar yapan araştırmacılar da düğümlerin işlevini anlayabilmek için böyle bir değişmezin keşfine ihtiyaç duyuyor. Geliştirilmesinin üzerinden geçen üç yılda AlphaFold, milyonlarca protein yapısı geliştirdi. Bu proteinlerin azımsanamayacak bir kısmının düğümlü yapıda olduğu biliniyor. Dolayısıyla düğümleri sınıflandırmak için kullanılabilecek bir topolojik değişmezin keşfi, proteinlerle ilgili bilimsel çalışmalarda yararlı olacaktır. Ayrıca düğümlerin önemli olduğu tek alan biyoloji de değil. Bilimsel çalışmalar, düğümlerin hidrojellerin termodinamik ve malzeme özelliklerini değiştirdiğini gösteriyor. Gelecekte yeni malzemeler tasarlayan araştırmacılar da topolojiden yararlanabilir. Yakın zamanlarda yapay zekânın düğümleri sınıflandırmada gösterdiği başarı, gelecek için umut vadediyor. ■

Kaynaklar

- Michieletto, Davide, “Spot the knot: using AI to untangle the topology of molecules”, *Physics World*, <https://physicsworld.com/a/spot-the-knot-using-ai-to-untangle-the-topology-of-molecules/>, 13 Ağustos 2024.
- Vandans, Olafs, ve ark., “Identifying knot types of polymer conformations by machine learning”, *Physical Review E*, Cilt 101, Makale No: 22502, 2020.
- Braghetto, Anna, ve ark., “Machine Learning Understands Knotted Polymers”, *Macromolecules*, Cilt 56, s. 2899, 2023.
- Sleiman, Joseph L., ve ark., “Geometric Predictors of Knotted and Linked Arcs”, *ACS Polymers Au*, Cilt 2, s. 341, 2022.
- Sleiman, Joseph L., ve ark., “Geometric learning of knot topology”, *Soft Matter*, Cilt 20, s. 71, 2024.
- Sulkowska, Joanna I., ve ark., “AlphaFold predicts novel human proteins with knots”, *Protein Science*, Cilt 32, Makale No: e4631, 2022.