

بیوانفورماتیک

مهدی صادقی*

زیست‌شناسی محاسباتی ارائه می‌دهد: بیوانفورماتیک عبارت است از تحقیق، توسعه یا کاربرد ابزارهای محاسباتی و نگرش‌هایی برای توسعه استفاده از داده‌های زیستی، پزشکی، رفتاری، و بهداشتی برای نگهداری، سازماندهی، طبقه‌بندی، تحلیل و مشاهده این داده‌ها.

زیست‌شناسی محاسباتی عبارت است از توسعه و کاربرد روش‌های نظری، مدل‌سازی ریاضی و تکنیک‌های شبیه‌سازی محاسباتی برای مطالعه سیستم‌های زیستی و رفتاری.

مرکز ملی اطلاعات زیست‌فناوری (آمریکا): بیوانفورماتیک رشته‌ای از علوم است که در آن زیست‌شناسی، علوم کامپیوتر، و تکنولوژی اطلاعات گرد هم آمده‌اند. سه زیررشته مهم در بیوانفورماتیک وجود دارد:

۱. ایجاد الگوریتم‌های جدید و شاخص‌های آماری که با آن بتوان رابطه بین اعضای مجموعه بزرگی از داده‌ها را سنجید؛

۲. تحلیل و تفسیر انواع مختلف داده‌ها شامل توالی اسیدهای آمینه و نوکلئوتیدها و ساختمان پروتئین‌ها؛

۳. توسعه و پیاده‌سازی ابزارهای مدیریت و دسترسی مؤثر به انواع مختلف داده‌ها و اطلاعات.

فرهنگ آکسفورد: «بیوانفورماتیک عبارت از زیست‌شناسی ادراکی در حیطه مولکولی و کاربرد تکنیک‌های انفورماتیکی (حاصل از رشته‌هایی نظیر ریاضیات کاربردی، علوم کامپیوتر و آمار) برای درک و سازماندهی اطلاعات موجود در این مولکول‌ها در مقیاس وسیع است. به طور خلاصه، بیوانفورماتیک سیستم مدیریت اطلاعات برای زیست‌شناسی مولکولی است که کاربردهای عملی وسیعی دارد.»

علاوه بر اینها می‌توان تعاریف دیگری نیز برای بیوانفورماتیک یافت. تفاوت در تعاریف بستگی دارد به حوزه مورد مطالعه و زمینه تخصصی شخصی که تعریف را ارائه می‌دهد. بیوانفورماتیک فقط کاربرد تکنیک‌های انفورماتیک و ریاضی در مطالعه موجود زنده نیست. پیوند بین این دو موضوع را می‌توان به این واقعیت نسبت داد که زیست‌شناسی ذاتاً یک

تلفیق علوم مختلف به منظور دستیابی به اهداف علمی خاص قدمتی طولانی دارد اما شکل‌گیری رسمی مقولاتی در علوم که میان رشته‌ای خوانده می‌شوند و افرادی با تخصص‌های گوناگون را به خود جذب کرده‌اند بیشتر در دهه‌های اخیر اتفاق افتاده است. این مقوله‌ها تعریف روشن و یا مرزبندی معینی با رشته‌های دخیل ندارند. ازجمله میان رشته‌ای‌های حوزه زیست‌شناسی می‌توان از مهندسی پزشکی، زیست‌ریاضی، زیست‌شناسی محاسباتی، زیست نانو تکنولوژی (یا نانو زیست تکنولوژی)، بیوشیمی، بیوفیزیک، و بیوانفورماتیک نام برد.

ارائه تعریفی برای هر یک از این رشته‌ها، که همپوشانی با رشته‌های دیگر هم نداشته باشد مشکل است. مثلاً در تعاریفی که معمولاً برای زیست‌ریاضی، زیست‌شناسی محاسباتی، و بیوانفورماتیک ارائه می‌شود نمی‌توان مرزبندی مشخصی را مشاهده کرد. زیست‌شناسی محاسباتی و ریاضی را می‌توان شامل کاربرد تکنیک‌های ریاضی، آمار، و علوم کامپیوتر در حل مسائل زیستی تعریف کرد، که اینها سابقه‌ای دیرینه‌تر از بیوانفورماتیک دارند. با این حال بیوانفورماتیک چه حوزه‌ای را در بر می‌گیرد و تعریف آن چیست؟ اصطلاح بیوانفورماتیک برای اولین بار در سال ۱۹۷۸ برای مطالعه پردازش اطلاعات در سیستم‌های زیستی به کار رفت. در طی سه دهه گذشته تعاریف متعددی برای بیوانفورماتیک ارائه شده است. به بعضی از این تعاریف در زیر اشاره می‌شود:

فرج تاکایا از انستیتو پاستور: «روش‌های ریاضی، آمار، و محاسباتی که به حل مسائل زیستی با استفاده از توالی‌های اسیدهای نوکلئیک و آمینه و سایر اطلاعات کمک می‌کند.»

ریچارد داربین از انستیتو سانگر: «من فکر نمی‌کنم بیوانفورماتیک شامل هر محاسبه زیستی باشد. مثلاً مدل‌سازی ریاضی، حتی وقتی که با مسائل مربوط به زیست‌شناسی ارتباط داشته باشد، بیوانفورماتیک نیست. به نظر من بیوانفورماتیک باید به مدیریت اطلاعات زیستی به‌ویژه اطلاعات ژنتیکی و مسائل متعاقب آن اطلاق شود.»

مؤسسه بهداشت ملی (آمریکا) تعریف دیگری از بیوانفورماتیک و

ساختمان پروتئین‌ها و جزئیات چگونگی برهم‌کنش محصولات را هم اضافه کرد و به این ترتیب می‌توان تصویری از مقدار و نوع اطلاعات در حال تولید داشت.

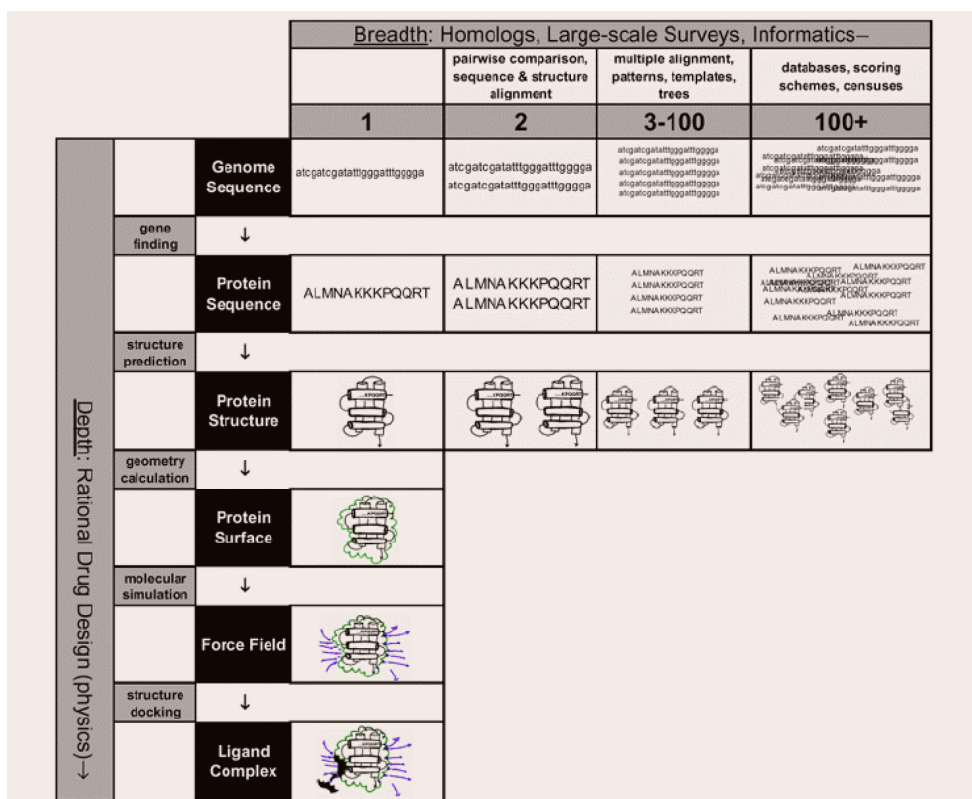
توسعه تکنیک‌های تعیین توالی ژنوم، در آینده‌ای نزدیک این امکان را فراهم می‌آورد که بتوان توالی ژنوم هر فرد را که مشخصه تمام ویژگی‌های ظاهری و باطنی مبتنی بر ژنتیک است تعیین کرد. در حال حاضر پروژه تعیین توالی ژنوم هزار نفر در حال تکمیل است که اطلاعات آن به تدریج در اختیار عموم قرار می‌گیرد. به این فهرست باید پروژه تعیین پلی‌مورفیسم و هاپلوتیپ صدها نفر از نژادهای مختلف را اضافه کرد که حاوی اطلاعات ذیقیمی درباره شباهت‌ها و تفاوت‌ها و تکامل نژادی انسان است. توسعه نرم‌افزاری و سخت‌افزاری امکان می‌دهد که این اطلاعات را در سطح کوچکی جا داد و به سرعت تحلیل کرد. این امر تأثیر غیر قابل پیش‌بینی بر سلامت و بهداشت خواهد گذاشت. درمان هر فرد بر مبنای خصوصیات ژنتیکی آن فرد طراحی و داروی مناسب تهیه و به کار برده خواهد شد. پروژه تعیین توالی ژنوم گونه‌های مختلف از باکتری‌ها تا گیاهان و جانوران منبع باارزشی از داده‌ها خواهد بود که در آینده نقشی اساسی در بیوتکنولوژی خواهد داشت.

مسائل اصلی در بحث سازماندهی و درک داده‌های زیستی در شکل ۱ خلاصه شده است. توسعه تکنیک‌های بیوانفورماتیک امکان گسترش تجزیه و تحلیل در دو بعد عمقی و عرضی را فراهم کرده است. بعد عمقی با محور عمودی در شکل نشان داده شده و رتوس یک نگرش ممکن به فرایند

فناوری اطلاعات است. فیزیولوژی و رفتار یک ارگانیسم عمدتاً توسط ژن‌های آن تعیین می‌شود که در ابتدایی‌ترین سطح می‌شود آنها را مخزنی دیجیتالی از اطلاعات تلقی کرد. موجود زنده نه فقط موضوع مطالعه با ابزارهای انفورماتیکی بلکه خود مجموعه‌ای از فرایندهای نگهداری و انتقال اطلاعات است.

در هر حال، فارغ از هر تعریفی که بتوان برای بیوانفورماتیک ارائه داد، واقعیت این است که به خصوص در دو دهه گذشته حجم وسیعی از داده‌های زیستی فراهم شده است که نگهداری، پردازش، و تفسیر آنها ممکن است تأثیرات عمیقی نه فقط در کیفیت زندگی، بلکه در درک ما از پیچیده‌ترین و عالی‌ترین سطح تجمع و کارکرد ماده یعنی حیات داشته باشد.

داده‌های زیستی با سرعت بی‌سابقه‌ای در حال افزایش است. برای مثال ذخیره اطلاعات توالی اسیدهای نوکلئیک تا پایان سال ۲۰۰۸ میلادی در GenBank محتوی حدود ۱۰۰,۰۰۰,۰۰۰ ورودی شامل حدود ۱۰۰,۰۰۰,۰۰۰ جفت باز و ذخیره اطلاعات توالی پروتئین‌ها در پایگاه داده‌های SWISS-PORT محتوی حدود ۴۰۵,۰۰۰ ورودی شامل ۱۴۷,۰۰۰,۰۰۰ اسید آمینه است. به طور متوسط، مقدار اطلاعات ذخیره شده در این پایگاه داده‌ها هر پانزده ماه دو برابر می‌شود. علاوه بر آن، از هنگام انتشار ژنوم influenza تا سال ۲۰۰۸، توالی کامل ژنوم ۵,۰۰۰ ارگانیسم به دست آمده که محتوی ۴۵۰ تا ۱۰۰,۰۰۰ ژن است. به این اطلاعات باید داده‌های حاصل از هزاران پروژه مرتبط بیان ژن‌ها، تعیین

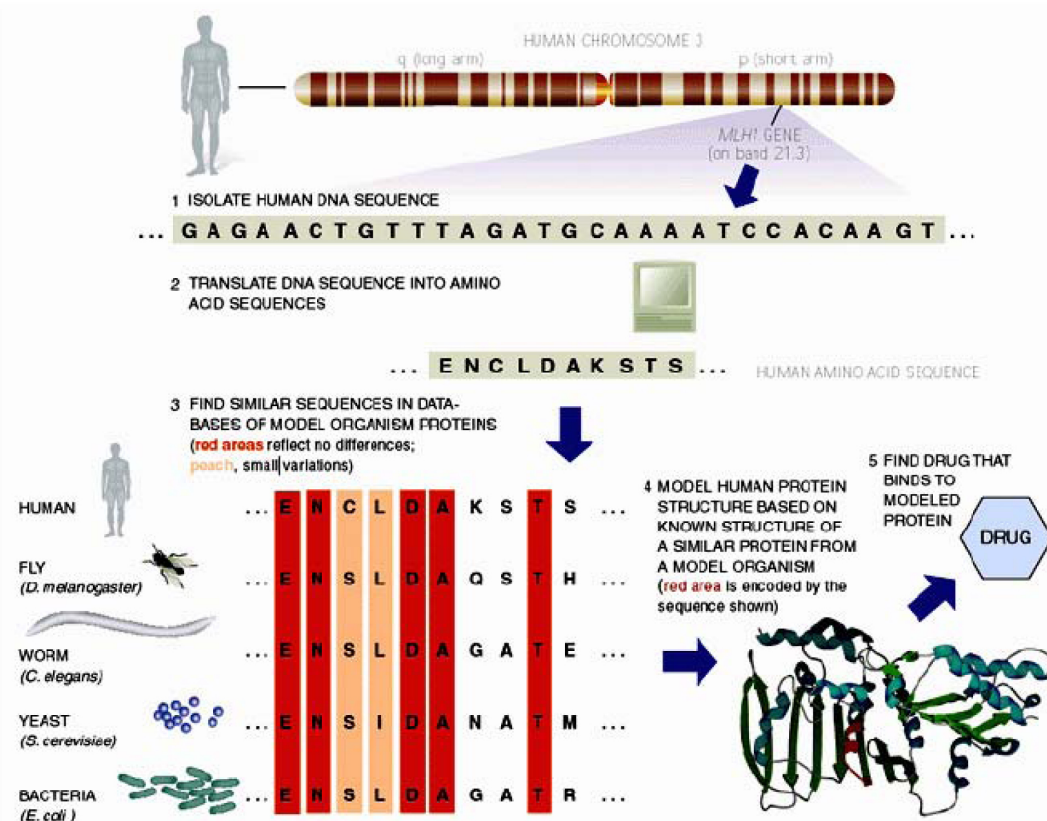


شکل ۱

یکی از کاربردهای اولیهٔ بیوانفورماتیک، کمک به طراحی منطقی داروها بوده است. در شکل ۱ نگرش عمومی مورد استفاده برای محصول یک ژن به عنوان مثالی از یک هدف دارویی خلاصه شده است. از طریق آنالیز لینکاژی و تشابه آن با ژن‌های مشابه در موش، نشان داده شده است که این ژن در سرطان رودهٔ بزرگ دخالت دارد. با اطلاع از توالی نوکلئوتیدی، توالی اسیدهای آمینه پروتئین کدگذاری شده را می‌توان با استفاده از نرم‌افزارهای ترجمه پیدا کرد. از روش‌های جستجوی توالی می‌توان برای پیدا کردن همولوگ‌ها در آرگانیزم‌های مدل استفاده کرد و براساس تشابه توالی می‌توان ساختمان پروتئین انسانی را با استفاده از ساختمان‌هایی که به طور تجربی تعیین شده‌اند مدل‌سازی کرد. در نهایت با الگوریتم‌های docking می‌توان مولکول‌هایی را که می‌توانند به ساختمان مدل متصل شوند طراحی و از سنجش‌های بیوشیمیایی برای آزمودن فعالیت زیستی آن بر روی پروتئین واقعی استفاده کرد.

در حال حاضر بیوانفورماتیک از مرحلهٔ مطالعهٔ یک ژن یا محصول به مطالعهٔ انبوه مولکول‌های زیستی در یک سلول یا آرگانیزم توسعه پیدا کرده است، اما این روند نیز به تدریج جای خود را به مطالعهٔ سیستم‌های پیچیده‌تر و شبکه‌های برهم‌کنش‌های این مولکول‌ها می‌دهد.

طراحی منطقی دارو را نمایش می‌دهد. هدف، گرفتن یک ژن و متعاقباً تجزیه و تحلیل آن است که موجب افزایش درک ما از پروتئینی می‌شود که کدگذاری می‌کند. با شروع از توالی ژن، می‌توان توالی پروتئین آن را به طور قطعی تعیین کرد و پس از آن، الگوریتم‌های پیشگویی را برای محاسبهٔ ساختمانی که پروتئین به خود می‌گیرد به کار برد. می‌توان با محاسبات هندسی سطح پروتئین را مشخص کرد و با شبیه‌سازی مولکولی میدان‌های نیروی اطراف مولکول را تعیین کرد. نهایتاً با استفاده از الگوریتم‌های docking، می‌توان لیگاندی را که به پروتئین متصل می‌شود طراحی کرد، که مثلاً می‌تواند دارویی باشد که عملکرد پروتئین را به طور اختصاصی تغییر می‌دهد. در عمل، انجام دقیق مراحل حد واسط هنوز مشکل است و بعضی از داده‌ها در ترکیب با روش‌های تجربی به دست می‌آید. هدف از بعد دوم یا عرضی، مقایسهٔ یک ژن با ژن‌های دیگر است. در ابتدا، از الگوریتم‌های ساده می‌توان برای مقایسهٔ توالی‌ها و ساختمان‌های یک جفت از پروتئین‌های مرتبط استفاده کرد. برای تعداد زیادتر پروتئین، الگوریتم‌های توسعه‌یافته را می‌توان برای ردیف‌سازی چندتایی به کار برد و طرح توالی یا الگوی ساختمانی را که معرف یک خانوادهٔ پروتئینی باشد استخراج کرد. با استفاده از این داده‌ها، ساخت درخت فیلوژنتیک برای تعقیب مسیر تکاملی پروتئین ممکن می‌شود. سرانجام، با داده‌های بیشتر، اطلاعات را باید در یک پایگاه بزرگ داده‌ها ذخیره کرد.



شکل ۲

مجازی، ظرفیت جدیدی را ایجاد می‌کند که محدودیت‌های آزمایشگاهی برای سنجش فرضیات جدید را ندارد. همراه با توسعه قابلیت‌های نرم‌افزاری و سخت‌افزاری و دانش انفورماتیک باید منتظر انقلابی در زیست‌شناسی و تغییر پارادایم‌های موجود باشیم که شاید تبعات آن در حوزه‌های مختلف علوم و فلسفه غیر قابل پیش‌بینی باشد.

* مهدی صادقی، پژوهشکده علوم کامپیوتر و پژوهشگاه ملی مهندسی ژنتیک و زیست‌فناوری.

آنالیز داده‌های زیستی در هر سطح و با هر پیچیدگی هدف اصلی بیوانفورماتیک است. این کار نیاز به استفاده روزافزون از ریاضیات، آمار، علوم کامپیوتر، و سایر علوم مرتبط دارد. اما نکته اساسی این است که تفسیر زیستی نتایج آنالیزها بر مبنای پارادایم غالب در زیست‌شناسی صورت می‌گیرد. این پارادایم که حاصل آزمایش‌های تجربی است در طی یک قرن گذشته براساس مشاهدات و فرضیاتی که سنجش آنها در لوله آزمایش یا در موجود زنده امکان داشته شکل گرفته است. حجم عظیم داده‌های خام موجود و در حال گسترش، امکانی فراتر از آنالیز آنها در پارادایم موجود را فراهم خواهد ساخت. در واقع بیوانفورماتیک به معنای زیست‌شناسی

آگهی درسی اختیاری

دربارهٔ اختیار

نیمسال اول ۱۳۸۷-۸۸

کاوه لاجوردی و مهدی نسرین

پژوهشکده فلسفه تحلیلی

فرض کنید شخص الف عمل ع را انجام داده است. آیا کار الف مختارانه بوده است؟ به نظر می‌رسد که جواب مثبت به این سؤال از جمله مستلزم این فرض باشد که الف می‌توانسته است ع را انجام ندهد. مثلاً اگر من دیروز ظهر مهرداد را به ضرب چاقو کشته‌ام، به نظر می‌رسد که فقط در صورتی این عمل‌ام با اختیار بوده است که، در زمان قتل، قادر بوده باشم به مهرداد چاقو زنم. اینکه (بخش معتناهی از) اعمال ما مختارانه هست یا نه البته مسئله‌ای بسیار قدیمی در فلسفه غرب و فلسفه اسلامی است. یکی از مسائل بحث اختیار این است که آیا فرض اختیار با علم خداوند به آنچه در آینده می‌کنیم سازگار هست یا نه، که این موضوع اصلی بحث ما نیست. از موضوعات اصلی این جلسات سازگار بودن اختیار با موجبیت‌گرایی (determinism) است — مطابق یکی از صورت‌بندی‌ها، موجبیت‌گرایی مدعی است که وضع هر ذره مادی در هر زمان t نتیجه قوانین طبیعی و وضع عالم در هر زمان قبلی t_0 است. پس مثلاً وضع ذرات تشکیل‌دهنده چاقو و دست من در ظهر دیروز نتیجه قوانین طبیعی و وضع عالم صد سال پیش از تولد من است. با فرض موجبیت‌گرایی، آیا جایی برای اختیار باقی می‌ماند؟ (به نظر می‌رسد که انکار موجبیت‌گرایی هم راه را برای اختیار باز نمی‌کند: اگر اینکه الف عمل ع را انجام می‌دهد نتیجه «تصادف» باشد، باز هم دشوار بتوان گفت که کار الف از روی اختیار بوده است.) در کنار نقد استدلال‌هایی برای ناسازگار بودن اختیار و موجبیت‌گرایی، به این هم خواهیم پرداخت که آیا اصولاً صدق موجبیت‌گرایی نتیجه‌ای در مورد مسئولیت (و اختیار) دارد یا نه. بحث به اخلاق هم مربوط می‌شود — از جمله به بررسی این مدعی رایج نیز می‌پردازیم که اگر الف نمی‌توانست کاری غیر از ع انجام دهد آنگاه الف در مورد ع مسئولیت اخلاقی ندارد.

جلسات از دوشنبه هفتم بهمن شروع می‌شود و یک هفته در میان در ساختمان نیاوران پژوهشگاه از ساعت چهار تا شش بعد از ظهر برگزار می‌شود. هشت متن فلسفه تحلیلی معاصر (نوشته رد ریک چیرلم، پیترو اینواگن، جی.جی.سی. سمارت، پتر سترلوسن، دیوید لوتیس، هری فرنکفورت، گیلین سترلوسن، تیشی آکانو) را بررسی خواهیم کرد؛ همه این مقالات در این مجموعه تجدید چاپ شده‌اند:

Garry Watson, ed., *Free Will*, Second Edition, Oxford, Oxford University Press, 2005.

نسخه‌ای از هر مقاله در کتابخانه پژوهشگاه موجود است. حضور برای همه آزاد است.