

شبکه های عصبی کوانتومی

علی روحانی، پیام خان تیموری

دانشکده برق دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده علوم کامپیوتر دانشگاه تهران

ali83online@yahoo.com

teymoori@sadi.ut.ac.ir

چکیده - روز به روز بر اهمیت محاسبات کوانتومی به علت دارا بودن توانایی و قدرت فراوان در انجام محاسبات و فراهم نمودن زمینه شناخت هر چه بیشتر کارکرد مغز و همچنین گسترش نانو کامپیوترها افزوده می شود. شبکه های عصبی کلاسیکی هم قدرت پردازش بالایی دارند که آن را مدیون پردازش موازی اطلاعات در خود هستند. از سویی دیگر مکانیک کوانتومی امکان فراهم نمودن پردازش موازی بسیار قوی تری را در ذات خود دارد، لذا به نظر می رسد که محاسبات کوانتومی می تواند در مقوله شبکه های عصبی کاملاً موثر واقع شود. هدف ما نیز در این مقاله ترکیب این دو مبحث با یکدیگر و بررسی ساختارهای ممکن شبکه های عصبی استوار بر محاسبات کوانتومی است.

کلیدواژه- شبکه های عصبی کلاسیکی، شبکه های عصبی کوانتومی، محاسبات کوانتومی، نظریه جهان های موازی.

1- مقدمه

شاید دو انگیزه اصلی برای بحث در زمینه شبکه های عصبی کوانتومی وجود داشته باشد. یکی از آنها ریشه در بحث هایی دارد که راجع به نقش فرآیند های کوانتومی در نحوه کارکرد مغز دارد و انگیزه دوم امکان ورود شبکه های عصبی کلاسیکی مصنوعی و نه بیولوژیکی به حوزه کوانتوم و بهره بردن از مزایای آن با ترکیب مناسب این دو فیلد با یکدیگر است.

البته ما در این مقاله توجه خود را به شبکه های عصبی مصنوعی معطوف می کنیم و به بررسی رهیافت های مختلف درک پردازش موازی کوانتومی خواهیم پرداخت و سپس یکی از این تفسیرهای مکانیک کوانتومی را به عنوان سنگ بنای ساخت و طرح ریزی شبکه های عصبی کوانتومی به کار خواهیم برد.

2- مفاهیم پایه ای در محاسبات کوانتومی

کارا ترین و متداول ترین مدل محاسبات کوانتومی مدل شبکه ای یا مدل Deutsch است که پایه ای ترین عنصر محاسباتی در آن بیت کوانتومی یا Qubit است. سیستم هایی با دو وضعیت گسسته و کاملاً مجزا از هم می توانند معرف یک بیت کوانتومی باشند. به عنوان مثال اتم هیدروژن یکی از همین سیستم ها است که دارای دو تراز مجزا از هم انرژی تحت عنوان حالت های پایه و برانگیخته است که می توانند معرف دو وضعیت صفر و یک باشند. سیستم های کوانتومی به کمک یک تابع موج ψ در فضای هیلبرت تعریف می شود، پایه ها در فضای هیلبرت بردارهای $|j_i\rangle$ هستند. سیستم کوانتومی به کمک رابطه زیر تعریف می شود.

$$|\Psi\rangle = \sum_{i,j \in \{0,1\}} a_{ij} |i\rangle_1 \otimes |j\rangle_2. \quad (4)$$

واحد های محاسباتی در مدل Deutsch شامل گیت های کوانتومی است که عملگرهای به اصطلاح Unitary هستند.

عملگرها در فضای هیلبرت بیانگر چگونگی تغییر یک تابع موج به دیگری هستند، که معمولاً با حروف بزرگ کلاه دار به شکل \hat{F} نشان داده می شوند و به شکل ماتریس های مربعی که در ماتریس های ستونی (بردار وضعیت کوانتومی) ضرب و بر روی آنها عمل می کنند بیان می شوند. شرط Unitary بودن برای تحقق محاسبات برگشت پذیر باید برقرار باشد، به عبارتی دیگر

$$\hat{F}^\dagger \hat{F} = \hat{F} \hat{F}^\dagger = \hat{I} \quad (5)$$

که \hat{I} عملگر یا ماتریس یکه و \hat{F}^\dagger ماتریس حاصل از ترانپوز (ماتریس) \hat{F} است.

3- تفاسیر گوناگون مکانیک کوانتومی

اگرچه امروزه ریاضیات مربوط به مکانیک کوانتومی کاملاً درک شده و مورد پذیرش قرار گرفته است، اما حقیقت فیزیکی مکانیک کوانتومی کاملاً مورد بحث است و توافقی کلی بر آن وجود ندارد، به عبارت واضح تر هیچ جای بحثی در روابط ریاضی مکانیک کوانتومی وجود ندارد، اما تفسیر و تعبیر آنچه که حقیقتاً اتفاق می افتد و چنین نتایجی را رقم می زند کاملاً جای بحث دارد. به تعاریف مختلفی که در رابطه با فرآیند های فیزیکی مکانیک کوانتومی وجود دارند در اصطلاح تفاسیر نظریه کوانتومی اطلاق می شود. یکی از این تعابیر "تفسیر کپنهاگ از مکانیک کوانتومی" است، تعبیر دیگر "تفسیر انتگرال مسیر فینمن" و دیگری "نظریه جهان های موازی" یا همان نظریه Many Universes (Multiverse) است که در سال 1957 توسط اورت مطرح شد.

درپایه گذاری و بیان مشابهات بین محاسبات کوانتومی و محاسبات عصبی اینکه کدامین تفسیر از مکانیک کوانتومی را به کار ببریم اهمیت بسیاری دارد. امروزه محققان

$$|y\rangle = \sum_i a_i |j_i\rangle \quad (1)$$

به این شکل که $|y\rangle$ در جمع آثاری از وضعیت های $|j_i\rangle$ به سر می برد، a_i ها ضرایب و در حالت کلی به صورت مختلط هستند [5].

هرگاه سیستم کوانتومی در جمع آثاری از پایه های خود باشد اصطلاحاً گفته می شود سیستم Coherent است، و هر گاه سیستم در اثر تعامل با محیط اطراف خود (فرآیند اندازه گیری) ترکیب خطی خود را از دست بدهد و تنها در وضعیت یکی از پایه ها قرار بگیرد اصطلاحاً گفته می شود که سیستم Decoherent است. از a_i ها که ضرایبی مختلط هستند تحت عنوان دامنه های احتمال یاد می شود زیرا که $|a_i|^2$ ها احتمال سقوط $|y\rangle$ به وضعیت $|j_i\rangle$ را بیان می کند و لذا این احتمالات باید تابع قانون احتمال باشند.

$$\sum_i |a_i|^2 = 1 \quad (2)$$

به عنوان مثال در مورد اتم هیدروژن که راجع به آن بحث شد، تابع موج سیستم در جمع آثاری از وضعیت های $|0\rangle$ و $|1\rangle$ به سر می برد. که هر یک از این وضعیت های صفر و یک اصطلاحاً پایه های محاسباتی نامیده می شوند.

$$|y\rangle = a|0\rangle + b|1\rangle \quad a^2 + b^2 = 1 \quad (3)$$

برای روشن تر شدن مطلب فرض کنید یک سیستم متشکل از دو بیت کوانتومی داریم، در نتیجه باید فضای هیلبرت ما دارای $N = 2^2$ پایه محاسباتی مجزا باشد.

فضای هیلبرت مرتبط با بیت کوانتومی اول را H_1 و متناسب با این انتخاب فضای مربوط به بیت کوانتومی دوم را H_2 می نامیم. پایه های محاسباتی برای این فضاها به قرار $\{|0\rangle_1, |1\rangle_1\}$ و $\{|0\rangle_2, |1\rangle_2\}$ هستند. اگر بیت کوانتومی 1 در وضعیت $|y\rangle_1 = a_{10}|0\rangle + a_{11}|1\rangle$ و بیت کوانتومی 2 در وضعیت $|y\rangle_2 = a_{20}|0\rangle + a_{21}|1\rangle$ باشند، آنگاه وضعیت سیستم متشکل از این دو بیت کوانتومی به صورت $|\Psi\rangle = |\Psi\rangle_1 \otimes |\Psi\rangle_2$ و به بیانی کامل تر در وضعیت زیر است [3].

می برد بدون اینکه وی از سایر نتایج ممکن آگاه باشد. به عنوان مثال هرگاه تاسی را می اندازیم همه حالات 1 تا 6 همزمان اتفاق می افتند، اما هر یک در جهانی از جهان های موازی و زمانی که ما نتیجه را مشاهده می کنیم (اندازه گیری) در یکی از این جهان ها قرار می گیریم و نتیجه مربوط به آن جهان را می بینیم در حالی که همه جواب های ممکن دیگر در جهان های دیگر به طور همزمان وجود دارند.

5- شبکه های عصبی

با نگاهی بر مراجع مختلف به نظر می رسد تعریفی عامه پسند از شبکه های عصبی (مصنوعی) وجود ندارد. اما در این بین اغلب محققان این زمینه در تعریفی از شبکه های عصبی که این شبکه ها را متشکل از واحد های پردازشی بسیار ساده که هریک مقداری جزئی حافظه محلی دارند اتفاق نظر دارند. نرون ها به کمک وزنه های ارتباطی به یکدیگر متصل می شوند.

برخی از شبکه های عصبی مدل هایی از شبکه های عصبی بیولوژیکی هستند و برخی دیگر اینگونه نیستند. تحقیقات و علاقه مندی به شبکه های عصبی از زمانی شروع شد که مغز به عنوان یک سیستم دینامیکی با ساختار موازی و پردازشگری کاملاً مغایر با پردازشگرهای متداول شناخته شد. لذا انگیزه های اولیه در زمینه شبکه های عصبی مصنوعی از تمایل به ساخت شبکه های مصنوعی که قادر به انجام محاسبات به اصطلاح هوشمند، همانند مغز انسان باشند نشأت گرفته است.

اگر مقاصد و اهداف یا به عبارتی دیگر تمامی اطلاعات مورد نیاز در اختیار نباشند نیاز به پروسه یادگیری داریم، در شبکه های عصبی، یادگیری به صورت تنظیم وزنه های ارتباطی رخ می نماید، به عبارتی شبکه های عصبی به کمک مثالها و الگوهای ورودی خروجی موجود یادگیری را انجام می دهند و اگر این فرآیند به خوبی انجام پذیرد این شبکه ها قادر خواهند بود جوابی مناسب برای الگوهایی که در فرآیند یادگیری دخیل نبودند ارائه دهند. شکل 1 مدلی از یک نرون را نشان می دهد، که در آن i بردار ورودی با n عضو (بیت)، o خروجی و w وزنه های ارتباطی مرتبط با ورودی هستند. تابع محرک هم می تواند از بین تابع های

گوناگون در سرتاسر جهان تفسیر های مختلفی را مبنای کار خود قرار می دهند اما به نظر می رسد که در مقوله محاسبات کوانتومی زیبا ترین تفسیر همان نظریه جهان های موازی باشد.

4- نظریه جهان های موازی

در این تفسیر از مکانیک کوانتومی تمامی عالم، فضای اطراف ما از کهکشان و ستاره ها گرفته تا اجسام کوچک، تمامی حقیقت را در بر نمی گیرند و تنها قسمتی کوچک از کل حقیقت فیزیکی را در بر می گیرند. به طوری که تعداد زیادی از مشابه جهان ما وجود دارند که هریک از آنها هم بخشی از حقیقت را در بر دارند [6].

به تعبیری جهان ما تقریبی از موجودیتی خود شمول در جهان های موازی است، به عبارتی دیگر جهان ما موجودیتی مجزا از جهان های دیگر است و جهان های دیگر تاثیری بر آن ندارند، این تقریب یا تعبیر از جهان، فیزیک کلاسیکی نامیده می شود و در نظریه محاسباتی از آن به محاسبات کلاسیکی و یا محاسبات تورینگ یاد می شود، و بدین معنی است که محاسبات در جهان ها متأثر از یکدیگر نیستند. در نقطه مقابل این نظریه، فیزیک کوانتومی جهان های موازی را مبنای کار خود قرار می دهد. به گونه ای که جهان ها از یکدیگر متأثر هستند.

در بسیاری از پدیده های فیزیکی و به خصوص پدیده های در ابعاد اتمی و میکروسکوپی، نظریه کلاسیکی از توصیف پدیده ها عاجز است زیرا در حقیقت اجسام فیزیکی که به یک جهان محدود می شوند، بسطی در سرتاسر جهان های موازی دارند، به عبارتی دیگر هر موجودیت در یک جهان همزاد هایی در جهان های دیگر دارد که این همزاد ها می توانند مستقل از هم عمل نمایند و بر یکدیگر اثر بگذارند. که این اثر گذاری تداخل کوانتومی نامیده می شود، که گواهی بر وجود حقیقت در ورای جهان ما هستند.

در این نظریه سقوط به یکی از پایه ها در اثر فرآیند اندازه گیری چیزی جز وهم نیست، زیرا که تابع موج ما در همه شرایط تابع معادلات شرودینگر است و در حقیقت زمانی که ما اندازه گیری انجام می دهیم تمامی جوابهای ممکن اتفاق می افتند اما هر یک در یکی از جهان های موازی و اندازه گیری مشاهده گر را به یکی از این جهان ها

از سوی دیگر اگر با هدف جلوگیری از تداخل الگوها از چندین شبکه هاپفیلد استفاده کنیم، در این صورت پردازش موازی الگوها را از دست می دهیم.

حال اگر از دیدگاه کوانتومی به مسئله نگاه کنیم می توانیم بر این مسئله اینگونه فائق آییم. تعدادی شبکه هاپفیلد (به تعداد الگوهای موجود) را به طور مجزا و هر یک را برای یک الگو آموزش می دهیم (ذخیره الگو)، هر شبکه و الگوی ذخیره شده در آن در یک جهان مستقل از دیگران وجود دارد، پس از اینکه هر شبکه در جهان خود (که مستقل از سایر جهانها است) الگوی مربوط به خود را ذخیره کرد، آنگاه جمع آثاری از این شبکه ها (الگوها) محاسبه و ساخته می شود و ما همه الگوها را یکجا داریم. به این رهیافت که در مرجع [8] نیز بدان اشاره شده است، شبکه متاثر ضعیف کوانتومی نامیده می شود. راه ساخته شدن چنین جمع آثاری نیز در [10] مورد بحث واقع شده است. در این متد هر یک از پایه های محاسباتی موجود در جمع آثار نقش یکی از الگوها را مستقل از دیگران دارد که فرایند اندازه گیری در نهایت موجب سقوط سیستم به یکی از این پایه ها (الگوها) خواهد شد.

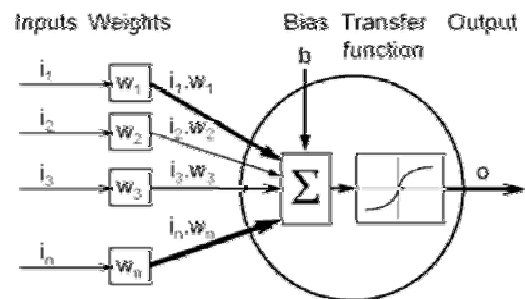
حال برای روشنتر شدن مفاهیم عنوان شده به ذکر مثالی می پردازیم.

به عنوان مثال فرض کنید مجموعه ای از سه شبکه هر یک با چهار بیت ورودی و یک بیت خروجی داریم، در نتیجه چهار وزنه ارتباطی خواهیم داشت که هر یک از ورودی ها را به خروجی مرتبط می سازند، ما این وزنه های ارتباطی را a_n, b_n, c_n, d_n نام گذاری می کنیم و داریم $n = 1, 2, 3$ ، به ازای شبکه های 1 تا 3. تمامی اطلاعات لازم در رابطه با شبکه ها در جدول شماره 1 جمع آوری شده است.

پس از اینکه شبکه ها به طور جداگانه آموزش دیدند، وزنه های ارتباطی با یکدیگر ترکیب و به صورت جمع آثار در می آیند، مثلاً وزنه ارتباطی اول که به بیت ورودی اول مرتبط است از جمع آثار $a1, a2, a3$ به دست می آید.

گوناگون انتخاب شود، به عنوان مثال SSGN می تواند نمونه خوبی برای تابع محرک باشد. در این صورت اگر مجموع حاصلضرب های ورودی در وزنه های مرتبطشان از حد آستانه b بزرگتر باشد خروجی نرون 1 و در غیر این صورت 0 خواهد شد [2].

$$O = f(W.i + b) \quad (6)$$



شکل 1: مدل یک نرون چند ورودی.

مدل عصبی ما هم اکنون نیاز به یک مکانیسم یادگیری دارد که معادله (7) می تواند رابطه مناسبی برای یادگیری باشد. این قانون موسوم به قانون یادگیری ویدرود- هوف است [2].

$$W(k+1) = W(k) + ae(k) i_k^T \quad (7)$$

تکرار الگوریتم فوق باعث تنظیم وزنه های ارتباطی می شود، که این وزنه های ارتباطی اطلاعات لازم برای حل مثال را در خود ذخیره دارند.

6- تحقق شبکه های عصبی کوانتومی

برای تحقق شبکه های عصبی کوانتومی می توان از مدل های گوناگون شبکه های عصبی بهره برد، ما نیز شبکه هاپفیلد را برای این منظور بر می گزینیم. (دلیل انتخاب این شبکه فراتر از حدود این مقاله است).

در شبکه های هاپفیلد کلاسیکی ما از یک شبکه برای ذخیره سازی چندین الگو استفاده می کنیم، اما این شبکه مشکلات خاص خود را داراست، از جمله اینکه برخی عناصر حافظه به علت تداخل با الگوهای دیگر ممکن است پایداری خود را از دست بدهند و این به نوبه خود گنجایش شبکه را کاهش می دهد، ضمناً شبکه هاپفیلد ممکن است به الگویی همگرا شود که جزو الگوهای ذخیره شده نباشد.

الگوی داده شده دارد برای پردازش اطلاعات انتخاب می شود.

7- برتری های شبکه های عصبی کوانتومی

با توجه به مطالب ارائه شده در مراجع [7,4,10,8] می توان گفت که شبکه های عصبی کوانتومی احتمالا دارای برتری های زیر در مقایسه با شبکه های عصبی مصنوعی کلاسیکی هستند.

- گنجایش ذخیره سازی اطلاعات به صورت توانی;
- عدم فراموشی الگوهای ذخیره شده به علت تداخل الگوهای ذخیره شده با یکدیگر;
- ابعاد کوچکتر (پیاده سازی در ابعاد اتمی);
- سرعت پردازش بالاتر اطلاعات;
- امکان پیاده سازی شبکه های تک لایه برای جداسازی الگوهایی که به صورت خطی از هم جداناپذیرند;

نتیجه گیری

در این مقاله ابتدا به معرفی ایده های ساده و در عین حال پایه ای از دو مبحث محاسبات کوانتومی و شبکه های عصبی مصنوعی پرداخته و سپس با بیان تعابیر گوناگون از مکانیک کوانتومی سعی نمودیم به بیان تشابهات این دو مقوله بپردازیم و پس از آن به طرح ریزی شبکه های عصبی کوانتومی پرداختیم و در خاتمه نیز به مزایای احتمالی این گونه شبکه ها نسبت به انواع کلاسیکی پرداختیم.

این زمینه مطالعاتی کاملا جدید است و به تعبیری در دوران طفولیت خود به سر می برد، لذا زمینه مطالعاتی بسیار گسترده ای در این فیلد وجود دارد.

رهیافت های دیگری نیز در این زمینه وجود دارد که فراتر از اهداف این مقاله بودند، یکی از آنها شبکه های شدیداً متاثر کوانتومی و دیگری رهیافتی است که بر مبنای نظریه گراف و جبر Rota پایه ریزی شده اند که خواننده می تواند با مراجعه به مراجع [8,9] اطلاعاتی راجع به آنها کسب کند.

Net	WEIGHT VALUES				TRAINING INPUT			
	1	2	3	4	1	2	3	4
1	a1	b1	c1	d1	0	0	0	1
2	a2	b2	c2	d2	1	0	1	0
3	a3	b3	c3	d3	1	1	1	1

جدول 1: بردارهای ورودی و وزنه های ارتباطی مربوط به 3 شبکه

هرگاه یک بیت ورودی به سیستم وارد می شود داخل جمع آثار وزنه های ارتباطی مرتبط با آن ورودی می شود. حال اگر بیت ورودی اول "1" باشد وارد جمع آثار a_2 و a_3 می شود زیرا بیت اول هردو این شبکه ها "1" است. به عبارتی دیگر a_2 و a_3 را می پذیرد و a_1 را نمی پذیرد. اگر پذیرفتن را با A و نپذیرفتن را با R نشان دهیم با فرض اینکه ورودی 1010 باشد خواهیم داشت.

1 ورودی بیت اول و داریم: Ra_1, Aa_2, Aa_3

0 ورودی بیت دوم و داریم: Ab_1, Ab_2, Rb_3

1 ورودی بیت سوم و داریم: Rc_1, Ac_2, Ac_3

0 ورودی بیت چهارم و داریم: Rd_1, Ad_2, Rd_3

مشاهده می کنیم که تنها جهانی که همه ورودی ها را پذیرفته است، جهان دوم (شبکه دوم) است پس این شبکه برای پردازش اطلاعات در این مثال به کار می رود. حال اگر الگوی ورودی به گونه ای باشد که دقیقا منطبق بر یکی از الگوها نباشد، یا به عبارتی دیگر جهانی وجود نداشته باشد که همه الگوها را بپذیرد، چه اتفاقی می افتد؟

فرض می کنیم که ورودی ما 1001 باشد که جزو هیچ یک از الگوهای ذخیره شده نیست. داریم:

1 ورودی بیت اول و داریم: Ra_1, Aa_2, Aa_3

0 ورودی بیت دوم و داریم: Ab_1, Ab_2, Rb_3

0 ورودی بیت سوم و داریم: Ac_1, Rc_2, Rc_3

1 ورودی بیت چهارم و داریم: Ad_1, Rd_2, Ad_3

از آنجایی که شبکه اول بیشتر از شبکه های دیگر ورودی ها را می پذیرد، یا به عبارتی دیگر بیشترین شباهت را به

- N.Kasabov,Ed.,Physica Verlag,pp. 213-235,2000.
- [8] Menneer, T, Narayanan, A, " Quantum Inspired Neural Networks", *NIPS 95, Denver, Colorado*, December 1996.
- [9] Altman, C, Pykacz, J, Zapatrin, R, R, "Superpositional Quantum Network Topologies", *arXiv:q-bio.NC/0311016* v3 20 Jul 2004.
- [10] Ventura, D, Martinez, T, "Initializing the amplitude distribution of quantum state" , submitted to *foundations of physics letters*, 1999.

سپاسگزاری

این مقاله بخشی از تحقیقات گروه در رابطه با شبکه های عصبی کوانتومی است. بر خود لازم می دانیم از جناب آقای دکتر منهای استاد محترم برق دانشگاه صنعتی امیرکبیر که هدایت گروه را برعهده دارند و همچنین مهندس علیرضا خان تیموری دانشجوی دکتری کامپیوتر دانشگاه صنعتی امیرکبیر از اعضای این گروه که کمک های بسیاری در تحقق این مقاله داشتند، نهایت تشکر را داشته باشیم.

نویسنده اول بر خود لازم می داند که از آقای مهندس ایمان اسماعیلی به خاطر کمک های شایانی که داشتند تشکر نماید.

مراجع

- [1] Nielsen. M.A & Chuang. I .L, *Quantum Computation and Quantum Information*, Cambridge University Press, 5th Printing, 2004.
- [2] Menhaj, .B., *Computational intelligence* Vol I, Amirkabir University press , Tehran, Iran,2000
- [3] Faber, J, Gilardi, G, A, "Quantum Models for Artificial Neural Networks", Brazil, 2003.
- [4] Ventura, D, "Artificial Associative Memory Using Quantum Processes", *Proceedings of International Conference on Computational Intelligence and Neuro-science*, Vol 2 , pp. 218-221 October 1998
- [5] Khanteymoori,A.R,Rohani,A, Khanteymoori ,P, "Introduction To Quantum Computing " ,*Amvaj-e-Bartar* ,Vol 19 ,2006.
- [6] Deutsch, D, "The Fabric of Reality", Alen Lane : The Penguin Press ,1997.
- [7] Ezhov , A,A, Ventura, D, "Quantum Neural Networks", *Future directions for intelligent information systems and information sciences*,