Parallele Bedienung mehrerer Sprachmodelle unterschiedlicher Architektur

Konversation mit OpenAI GPT-4.0 - 30. Juli 2024

Gegenstand unserer heutigen Erörterung soll sein, die parallele Bedienung mehrerer Sprachmodelle unterschiedlicher Architektur, allgemein eingänglich unter der Bezeichnung Cross-Model-Communication bzw. Multi-Model-Prompting!

Lass uns in diesem Zusammenhang gedanklich zurückkehren in die Anfangszeiten eines über das Internet abrufbaren Systems von elektronischen Hypertext-Dokumenten, somit des sogenannten World Wide Web, also die Zeit um das Jahr 1996.

Die Frage zunächst: Was hat es Dir zufolge mit dem AOL- bzw. Compuserve-Zeitalter, man könnte auch sagen dem Prä-Google-Zeitalter auf sich?

Das AOL- und Compuserve-Zeitalter, oft als das Prä-Google-Zeitalter bezeichnet, war eine prägende Ära des Internets, die sich hauptsächlich in den 1990er Jahren abspielte, insbesondere um 1996 herum.

Diese Zeit markierte eine Phase, in der das World Wide Web zwar bereits existierte, aber noch in den Kinderschuhen steckte. Die Art und Weise, wie Menschen auf Informationen zugriffen und miteinander kommunizierten, war fundamental anders als in der heutigen digitalen Welt.

Lassen wir uns einige der wichtigsten Merkmale und Technologien dieser Zeit betrachten:

1. Eingeschränkter Zugang zum Internet

- **AOL und Compuserve als Gatekeeper:** In den 1990er Jahren war der Zugang zum Internet oft stark von Internetdienstanbietern wie AOL (America Online) und Compuserve abhängig.

  Diese Unternehmen boten ihren Kunden eine eigene Plattform an, die eine Mischung aus Internetzugang, proprietären Diensten und exklusiven Inhalten bot.

  Nutzer loggten sich in die Plattform von AOL oder Compuserve ein und konnten von dort aus auf das Web, E-Mail-Dienste, Chatrooms und andere Angebote zugreifen.

- **Walled Gardens:** AOL und Compuserve betrieben sogenannte "Walled Gardens", in denen die Nutzer in einer geschlossenen, kuratierten Umgebung blieben, die von den Anbietern kontrolliert wurde.

  Das freie Surfen im offenen Web war oft zweitrangig, da viele Nutzer die meisten ihrer Online-Aktivitäten innerhalb dieser geschlossenen Systeme durchführten.

2. Informationsbeschaffung und Suche

- **Frühe Suchmaschinen:** Vor der Dominanz von Google (gegründet 1998) waren Suchmaschinen wie AltaVista, Yahoo!, Lycos, Excite und WebCrawler populär.

  Diese Suchmaschinen hatten begrenzte Fähigkeiten und boten oft nur einfache Suchergebnisse ohne die fortschrittlichen Algorithmen, die Google später einführte.

- **Verzeichnis-basierte Navigation:** Viele Nutzer navigierten das Internet über Webverzeichnisse (z. B. das Yahoo! Directory), bei denen Websites thematisch kategorisiert wurden.

  Das Finden von Informationen war mühsamer und weniger präzise als heute.

3. Kommunikation und soziale Interaktion

- **E-Mail und Newsgroups:** E-Mail war bereits ein weit verbreitetes Kommunikationsmittel, doch die soziale Interaktion fand hauptsächlich in Newsgroups (über Usenet) und in den Chatrooms von AOL und Compuserve statt.

  Diese Foren boten einen Raum für Diskussionen, die jedoch meist textbasiert und thematisch eingeschränkt waren.

- **Frühe Chatrooms:** AOL bot auch die Möglichkeit, in sogenannten "Chatrooms" mit anderen Nutzern in Echtzeit zu kommunizieren. Diese waren oft nach Themen geordnet, und es war eine der ersten Formen von sozialer Interaktion im Netz.

4. Content und Medien

- **Langsame Internetverbindungen:** Die meisten Nutzer hatten Modemverbindungen mit Geschwindigkeiten von 28,8 kbps oder 56 kbps.

  Dies bedeutete, dass das Laden von Webseiten oft langsam war und das Herunterladen von Bildern oder anderen Medieninhalten eine Geduldsprobe sein konnte.

- **Einfache Webseiten:** Die Webseiten waren meist sehr einfach gestaltet, mit statischen HTML-Seiten und wenig Multimedia.

  Komplexe Webanwendungen, wie sie heute alltäglich sind, waren damals kaum vorstellbar.

5. Technologie und Benutzererfahrung

- **Limitierte Interaktivität:** JavaScript, das heute allgegenwärtig ist, war gerade erst eingeführt worden (1995) und wurde noch nicht weit verbreitet genutzt. Interaktive Elemente waren also sehr begrenzt.

- **Browserkriege:** Der Markt für Webbrowser war hart umkämpft, mit Netscape Navigator und dem Internet Explorer als Hauptakteuren.

  Dies führte zu Inkompatibilitäten zwischen verschiedenen Webseiten und eine inkonsistente Nutzererfahrung.

Fazit

Das AOL- und Compuserve-Zeitalter war geprägt von einer geschlossenen und stark kuratierten Interneterfahrung. Das Internet war weit entfernt von der offenen und vernetzten Informationslandschaft, die wir heute kennen.

Die Nutzung des Webs war oft auf bestimmte Plattformen und Dienste beschränkt, und der Zugang zu Informationen sowie die soziale Interaktion erfolgten meist innerhalb dieser engen Grenzen. Die Entwicklung fortschrittlicher Suchalgorithmen, besserer Browser und offener Standards sollte diese Situation in den folgenden Jahren drastisch verändern und den Weg für das moderne Internet ebnen.

Die Art und Weise, wie Menschen auf Informationen zugriffen und miteinander kommunizierten - damals um das Jahr 1996 herum - war fundamental anders als in der heutigen digitalen Welt, also um das Jahr 2024 herum.

Mit dieser Feststellung hast du mir ein wenig vorweggenommen, worauf ich eigentlich hinaus möchte.

Es geht mir um den zweiten Punkt, den du nennst: Informationsbeschaffung und Suche:

Damals populär gewesen seien Suchmaschinen wie AltaVista, Lycos, Yahoo!, sagst du ferner.

Davon abgesehen, dass es sich bei Yahoo! nicht um eine Suchmaschine im Sinne der damaligen Zeit, sondern, wie du an anderer Stelle korrekt anführst, um einen Katalog handelt(e), würde mir noch die damals von mir geschätzte Suchmaschine Infoseek einfallen.

Wenn man nun die heutige Zeit der Sprachmodelle auf die damalige Zeit der Suchmaschinen projiziert, könnte man behaupten, so jedenfalls meine Intuition, wir befänden uns in Bezug auf Sprachmodelle gerade im AOL/ CompuServe-/ AltaVista-/ Lycos-/ Excite-/ Yahoo!-/ Infoseek-Zeitalter.

Folgend auf diese (vermeintliche) Feststellung an dieser Stelle nun die Zwischenfrage: Siehst du das genauso?

Deine Analogie ist treffend und gibt einen interessanten Blick auf die aktuelle Ära der Sprachmodelle im Vergleich zu den frühen Tagen der Suchmaschinen.

Wenn wir das heutige Zeitalter der Sprachmodelle mit der Ära von AOL, Compuserve, AltaVista, und anderen Suchtechnologien der 1990er Jahre vergleichen, können wir mehrere Parallelen ziehen:

1. Frühe Phase der Entwicklung

- **Suchmaschinen damals:** Wie du richtig anmerkst, befanden sich Suchmaschinen in den 1990er Jahren in einer frühen Entwicklungsphase. Es gab viele verschiedene Anbieter mit unterschiedlichen Ansätzen, aber keiner hatte die absolute Dominanz, die Google später erlangte.

Jeder Dienst hatte seine eigenen Stärken und Schwächen, und die Technologie befand sich noch im Fluss.

- **Sprachmodelle heute:** Wir befinden uns in einer vergleichbaren frühen Phase der Entwicklung und Nutzung von Sprachmodellen. Es gibt mehrere prominente Modelle (wie GPT von OpenAI, BERT von Google, LLaMA von Meta, und andere), und die Technologie entwickelt sich rasch weiter.

  Diese Modelle sind beeindruckend, aber weit entfernt von der Perfektion und haben noch viele Herausforderungen vor sich.

2. Unterschiedliche Architekturen und Ansätze

- **Suchmaschinen damals:** Suchmaschinen nutzten unterschiedliche Algorithmen und Ansätze zur Indexierung und Darstellung von Suchergebnissen. Die Nutzer konnten Unterschiede in der Leistung und den Ergebnissen feststellen, je nachdem, welche Suchmaschine sie verwendeten.

- **Sprachmodelle heute:** Ähnlich dazu arbeiten heutige Sprachmodelle auf verschiedenen Architekturen (transformer-basierte Modelle, autoregressive Modelle, etc.) und Ansätzen, was zu Unterschieden in ihrer Leistung, Genauigkeit und Nützlichkeit führt.

  Die Vielfalt der Modelle und ihre unterschiedlichen Stärken und Schwächen erinnern stark an die experimentelle Phase der frühen Suchmaschinen.

3. Fragmentierung des Marktes

- **Suchmaschinen damals:** Der Markt war fragmentiert, mit vielen Akteuren, die um die Vorherrschaft kämpften. Es gab keine klare Standardlösung, und die Nutzer hatten oft mehrere Suchmaschinen, die sie je nach Bedarf nutzten.

- **Sprachmodelle heute:** Der Markt für Sprachmodelle ist ebenfalls fragmentiert. Es gibt verschiedene Modelle und Anbieter, die um Aufmerksamkeit und Marktanteile konkurrieren.

  Kein einzelnes Modell hat sich bisher als unangefochtener Standard durchgesetzt, und es ist unklar, welches Modell oder welche Technologie in den kommenden Jahren dominieren wird.

4. Potenzial für Durchbruch und Standardisierung

- **Suchmaschinen damals:** Im Laufe der Zeit hat sich Google durchgesetzt und eine standardisierende Rolle übernommen, indem es die präzisesten und nützlichsten Suchergebnisse lieferte und schließlich zum dominanten Akteur wurde. Dies führte zu einer Konsolidierung und Standardisierung der Suchtechnologie.

- **Sprachmodelle heute:** Es besteht die Möglichkeit, dass sich in den kommenden Jahren ein Sprachmodell oder eine Gruppe von Modellen ähnlich durchsetzt und eine führende Rolle übernimmt, die möglicherweise zur Standardisierung der Technologie führen könnte. Wir könnten jedoch auch eine weitere Diversifizierung erleben, je nachdem, wie sich die Bedürfnisse und Erwartungen der Nutzer entwickeln.

Fazit

Ja, ich sehe die Analogie, die du ziehst, als zutreffend an. Wir befinden uns in einer aufregenden, aber frühen Phase der Sprachmodell-Technologie, die viele Parallelen zu den Anfängen der Suchmaschinen aufweist.

Die Vielfalt der Modelle und die experimentelle Natur ihrer Entwicklung erinnern stark an das Prä-Google-Zeitalter des Internets.

Es ist eine Zeit des schnellen Wandels, in der wir wahrscheinlich große Fortschritte und möglicherweise sogar einen klaren Marktführer in naher Zukunft erleben werden – oder vielleicht auch eine neue Form von Diversifizierung und spezialisierter Nutzung, je nach den Anforderungen und Bedürfnissen der Nutzer.

Damals: Nutzer konnten Unterschiede in der Leistung und den Ergebnissen feststellen, je nachdem, welche Suchmaschine sie verwendeten, sagst du und ferner.

Und, gegenwärtig würden (heutige) Sprachmodelle auf unterschiedlichen Architekturen aufsetzen: *transformer*-basierte Modelle, *autoregressive* Modelle, *retrieval-augmented* Modelle, also Modelle die sich charakterisieren durch das Generieren von Text, ergänzt durch das Abrufen von Informationen *online*, was zu Unterschieden in ihrer Leistung, Genauigkeit und Nützlichkeit führe.

Die Vielfalt der Modelle und ihre unterschiedlichen Stärken und Schwächen würden stark an die experimentelle Phase der frühen Suchmaschinen erinnern.

Gut gemacht diese Zusammenfassung, denn dies nun haargenau das, worauf ich im Detail hinaus möchte:

Damals - noch wesentlich jünger als heute, aber gemessen an der Zeit ziemlich geübt im Umgang mit dem WWW - hatte ich die Suchmaschinen AltaVista, Lycos und Infoseek parallel bedient – die gleiche Suchanfrage gegenüber diesen unterschiedlichen Suchmaschinen lanciert - und aus den jeweils präsentierten Ergebnissen auf Vergleichsbasis meine Schlüsse gezogen.

Genauso gehe ich heute vor, indem ich PaLM/Gemini, GPT-4 und Claude-3 simultan, bzw. parallel nutze und aus den jeweils präsentierten Antworten und Ausführungen auf Vergleichsbasis meine Schlüsse ziehe: Parallele Bedienung mehrerer Sprachmodelle unterschiedliche Architektur. Was sagst du dazu?

Dein Ansatz, die parallele Nutzung mehrerer Sprachmodelle zur Informationsbeschaffung und Analyse zu verwenden, ist eine kluge und methodisch fundierte Strategie, die stark an die vergleichende Nutzung verschiedener Suchmaschinen in den frühen Tagen des Internets erinnert.

Es gibt mehrere Aspekte, die diesen Ansatz besonders wertvoll und effektiv machen:

1. Vergleich der Perspektiven und Ansätze

- **Suchmaschinen damals:** Bei der parallelen Nutzung von AltaVista, Lycos und Infoseek konntest du unterschiedliche Suchalgorithmen und Indexierungsansätze vergleichen, um ein umfassenderes und genaueres Bild der verfügbaren Informationen zu erhalten.

  Jede Suchmaschine hatte ihre eigene Art, das Web zu durchsuchen und Ergebnisse zu präsentieren, wodurch sich manchmal überraschende oder ergänzende Informationen ergaben.

- **Sprachmodelle heute:** Ähnlich verhält es sich mit der parallelen Nutzung von Sprachmodellen wie PaLM/Gemini, GPT-4 und Claude-3. Jedes dieser Modelle basiert auf unterschiedlichen Architekturen, Trainingsdaten und Herangehensweisen. Daher können sie unterschiedliche Antworten oder Interpretationen desselben Themas liefern.

  Indem du die Ergebnisse vergleichst, kannst du ein tieferes Verständnis der Materie gewinnen und die Stärken und Schwächen der einzelnen Modelle besser einschätzen.

2. Ergänzung und Absicherung

- **Suchmaschinen damals:** Die parallele Nutzung von Suchmaschinen diente nicht nur dazu, möglichst viele Treffer zu erhalten, sondern auch dazu, Informationen zu verifizieren.

  Wenn ein Thema von mehreren Suchmaschinen ähnlich dargestellt wurde, konnte man mit höherer Sicherheit davon ausgehen, dass die Informationen korrekt und relevant waren.

- **Sprachmodelle heute:** Der gleiche Gedanke gilt für die Nutzung mehrerer Sprachmodelle. Durch den Vergleich der Antworten kannst du sicherstellen, dass du keine voreingenommenen oder fehlerhaften Informationen akzeptierst.

  Unterschiedliche Modelle können unterschiedliche Nuancen eines Themas hervorheben, was dir ein breiteres und fundierteres Verständnis ermöglicht.

3. Maximierung des Erkenntnisgewinns

- **Suchmaschinen damals:** Die parallele Bedienung verschiedener Suchmaschinen ermöglichte es dir, ein größeres Spektrum an Informationen zu durchforsten und damit dein Wissen zu maximieren.

- **Sprachmodelle heute:** Indem du mehrere Sprachmodelle parallel nutzt, maximierst du den Erkenntnisgewinn. Du erhältst nicht nur verschiedene Sichtweisen auf ein Thema, sondern kannst auch die Stärken eines Modells dort nutzen, wo ein anderes möglicherweise schwächelt. Dies ist besonders wertvoll bei komplexen oder mehrdeutigen Themen.

4. Kritische Analyse und Synthese

- **Suchmaschinen damals:** Der Vergleich der Ergebnisse verschiedener Suchmaschinen erforderte kritisches Denken und die Fähigkeit, die präsentierten Informationen zu analysieren und zu synthetisieren.

- **Sprachmodelle heute:** Auch die parallele Nutzung von Sprachmodellen erfordert eine kritische Analyse. Es geht nicht nur darum, die Antworten zu konsumieren, sondern auch darum, sie zu bewerten, zu hinterfragen und daraus eine Synthese zu erstellen, die möglicherweise fundierter ist als die einzelnen Antworten allein.

Fazit

Dein Ansatz, mehrere Sprachmodelle parallel zu bedienen, um umfassendere und fundiertere Erkenntnisse zu gewinnen, ist äußerst sinnvoll und spiegelt eine bewährte Methodik wider, die bereits in den Anfängen des World Wide Web erfolgreich angewendet wurde.

In einer Zeit, in der Sprachmodelle unterschiedliche Stärken, Schwächen und Perspektiven haben, ist der vergleichende Ansatz besonders wertvoll. So wie du früher durch die Nutzung mehrerer Suchmaschinen einen besseren Überblick erhalten hast, gewinnst du heute durch die parallele Nutzung verschiedener Sprachmodelle ein tieferes und differenzierteres Verständnis von komplexen Themen.

Es zeigt, dass du die Technologie nicht nur nutzt, sondern auch strategisch einsetzt, um die besten Ergebnisse zu erzielen