

Lehrstuhl für Softwaretechnik
Prof. Dr. Stephan Diehl
Universität Trier
Fachbereich IV – Informatik

Vergleich von Graphlayoutverfahren mit Hilfe von Eyetracking und Task-orientierter Analyse

Diplomarbeit

vorgelegt von

Markus Schmitt

Matrikel-Nummer: 657712

Neuhütten im Februar 2008

Erklärung der Selbstständigkeit

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die Zitate deutlich kenntlich gemacht zu haben. Die Arbeit enthält nach bestem Wissen kein Material, das bereits veröffentlicht oder von anderen Personen geschrieben worden ist.

Neuhütten, den 15.02.2008

Markus Schmitt

Danksagung

Diese Arbeit bietet mir nicht nur die Möglichkeit, neue wissenschaftliche Pfade zu beschreiten, sondern gibt mir auch die Gelegenheit, den Menschen zu danken, die zum Erfolg dieser Arbeit (und meines Studiums) beigetragen haben.

Bedanken möchte ich mich bei Prof. Dr. Stephan Diehl, der durch die Vergabe dieses eher ungewöhnlichen Themas am Lehrstuhl für Softwaretechnik den Weg für diese Arbeit ebnete. Es ist in meinen Augen keine Selbstverständlichkeit, eine Diplomarbeit genau auf die Interessen und Fähigkeiten eines Diplomanden abzustimmen. Dies weiß ich sehr zu schätzen.

Einen großen Einfluss auf diese Arbeit hatte auch Mathias Pohl, der nicht nur in langen und intensiven Treffen mit Prof. Diehl und mir die Rahmenbedingungen absteckte und den Grundaufbau der Evaluation entwarf, sondern stets ein kompetenter Ansprechpartner und Berater bei allen fachlichen und gestalterischen Fragen war. Besonders seine \LaTeX -Erfahrung kam mir in den letzten Wochen der Entstehung dieser Arbeit zugute. Ein herzliches Dankeschön dafür.

Mein Dank gilt auch Prof. Dr. Stefan Näher für die Übernahme der Zweitprüfung dieser Arbeit mit einer ebenfalls für ihn sicherlich nicht alltäglichen Thematik, den Teilnehmern der Vorstudie, der Hauptevaluation und der Zusatzstudie, der Abteilung Kartographie für die Bereitstellung der technischen Mittel (hier besonders Kerstin Servatius als kompetente Ansprechpartnerin), sowie Felix Bott, der mir während der Evaluationsphase eine große Hilfe war.

Last but not least gilt mein ganz besonderer Dank den Menschen, ohne die ein erfolgreiches Studium kaum möglich gewesen wäre: Für ihre bedingungslose Unterstützung in den vergangenen Jahren danke ich meinen Eltern und meiner Freundin Christina, die nicht nur ungezählte Stunden mit dem Korrigieren meiner Arbeit verbrachte, sondern mich stets auch in den Phasen meines Studiums motivierte, in denen ich einen kleinen Anschub benötigte.

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation	3
1.2	Übersicht	7
2	Theoretische Grundlagen	9
2.1	Graphentheoretische Grundlagen	9
2.2	Graphlayouts	13
2.2.1	Das kräftebasierte Layoutverfahren	13
2.2.2	Das orthogonale Layoutverfahren	18
2.2.3	Das hierarchische Layoutverfahren	23
2.2.4	DGD – Dynamic Graph Drawing	27
2.3	Eyetracking – Grundlagen und Technik	29
2.4	Die verwendeten Tasks	31
2.4.1	Task 1 – Existenz eines Knotens	31
2.4.2	Task 2 – Wegeproblem	32
2.4.3	Task 3 – Existenz eines Teilgraphen	33
2.4.4	Task 4 – Cliquen-Problem	35
2.4.5	Task 5 – Maximaler Grad	36
3	Durchführung und Auswertung der Studie	38
3.1	Versuchsaufbau und -rahmenbedingungen	38
3.1.1	Probanden	39
3.1.2	Design der Untersuchung	40
3.1.3	Technische Ausstattung	44
3.1.4	Ablauf der Studie	45
3.2	Die Vorstudie	48
3.3	Quantitative Analyse	62
3.3.1	Auswertung der Korrektheit der Antworten	63
3.3.2	Auswertung der Bearbeitungszeiten	72
3.3.3	Weitere Auswertungen	78
3.4	Qualitative Analyse	85

3.4.1	Heatmap-Analyse	86
4	Zusammenfassung und Ausblick	112
4.1	Zusammenfassung	112
4.2	Beurteilung und Ausblick	115
	Literaturverzeichnis	117

1 Einführung

Ein Bauer will einen Fluss überqueren, dabei möchte er einen Wolf, eine Ziege und einen Kohlkopf mitnehmen. Sein Boot ist gerade groß genug, um ihn selbst und ein weiteres Objekt zu transportieren. Er kann also immer nur eines der drei Objekte mit sich auf die andere Seite bringen. Lässt er den Wolf mit der Ziege alleine, so wird die Ziege gefressen werden. Lässt er die Ziege mit dem Kohlkopf zurück, dann wird diese den Kohlkopf verzehren.

Gibt es für den Bauern eine Möglichkeit den Fluss zu überqueren, ohne dass die Ziege oder der Kohlkopf gefressen werden? Und wenn ja, wie viele Fahrten benötigt der Bauer mindestens?

Diese Denkaufgabe ist über 1200 Jahre alt und stammt aus der Feder des Mönches und Abts des Klosters St. Martin in Tours Alkuin. Möchte man eine Lösung dieser Aufgabe auf eine systematische Art und Weise suchen, so findet man sich unvermittelt in der Welt der Graphentheorie wieder, denn der einfachste Weg die Aufgabe zu lösen, ist die Modellierung der Situation mit Hilfe eines Graphen, der die möglichen und unmöglichen Lösungen gegenüberstellt.

Ist die Graphentheorie also ein Relikt aus einer längst vergangenen Zeit, dass für unsere heutige moderne Welt an Bedeutung verloren hat? Ganz im Gegenteil. Diese Disziplin der Mathematik erhielt sehr früh ihren Einzug in die Informatik und spielt zum Beispiel bei der Informationsvisualisierung, einem noch recht jungen Forschungsgebiet der Informatik, eine tragende Rolle. Man erkannte sehr schnell, dass die Informationsvisualisierung ein sehr weitreichendes und interdisziplinäres Forschungsfeld ist, das mit der zunehmenden Komplexität der heutigen Welt permanent an Bedeutung gewinnt. Aber Graphen stellen in der Informatik nicht nur in der Informations- und Softwarevisualisierung eine entscheidende Stütze dar, sondern sie finden in allen Bereichen der Informatik ihre Verwendung als Modellierwerkzeug, sei es bei der Automatentheorie, der Berechenbarkeits- oder der Komplexitätstheorie, bei der Modellierung von Datenbanken, Netzwerken oder bei Betriebs- und Dateisystemen.

Man könnte sagen: Wir sind heutzutage überflutet von Daten. Egal ob die Verantwortlichen eines Unternehmens am Ende des Jahres Bilanz über Gewinne und Verluste ziehen, um daraus eine künftige Firmenpolitik abzuleiten, ob ein Kaufinteressent sich einen Vergleichstest der Automobile betrachtet, die er in die engere Wahl genommen hat oder ob Statistiken einer Fußball-Bundesligasaison in eine für den Fernsehschauer leicht verständliche Form gebracht werden sollen, überall müssen sich die betreffenden Personen Gedanken über eine bestmögliche Darstellung der nackten Daten machen.

Unglücklicherweise liegt eben gerade hier das Problem: Meist liegen die Daten als ein unüberschaubares Meer von Zahlen und Text vor, so dass sich das Verstehen und Darstellen dieser Werte als sehr schwierig und problematisch darstellt. Was ist also nötig, um die Übersicht und letztendlich auch die Herrschaft über die großen Datenmengen nicht zu verlieren? Die Antworten, die die Informatik hier geben kann sind eine bessere Organisation der Daten, die eine sinnvolle Struktur anbietet und das Verleihen einer mit Semantik angereicherten Repräsentation, die dem Betrachter der Daten eine schnelle Erfassung und ein umfassendes Verstehen der Bedeutung erleichtert.

Die Aktualität dieses Themas lässt sich sehr gut an der hohen Forschungsaktivität in diesem Bereich der Informatik erkennen. Es werden regelmäßig neue Tools und Strategien veröffentlicht, die bestimmte Daten besonders gut und verständlich darstellen sollen. Auch jährlich stattfindende Konferenzen wie InfoVis [inf] und GD (Graph Drawing) [gdc] geben einen Überblick über den aktuellen Stand der Forschung. Neue Wege der Informationsvisualisierung werden permanent besprochen, hier gilt es allerdings das vorrangige Ziel, nämlich die Übersichtlichkeit und Verständlichkeit für einen menschlichen Betrachter, permanent und immer wieder zu prüfen und zu optimieren. Diese „menschliche Komponente“ darf keinesfalls gegenüber den informationstechnischen Fortschritten an Bedeutung verlieren, denn der Adressat einer Visualisierung von Daten ist immer der Mensch, ein Computer kann wesentlich leichter mit einer Ansammlung von Rohdaten zurechtkommen.

Einen hohen Stellenwert hat die Informationsvisualisierung auch in Forschungsgebieten erreicht, die auf den ersten Blick wenig mit der Informatik in Verbindung stehen. So benötigt man sowohl in der Sozialforschung als auch in der Psychologie die Werkzeuge, die in der Informatik entwickelt wurden, zum Beispiel bei der Analyse von sozialen Netzwerken. Im Zuge des „Web 2.0“-Booms schossen Internet - Plattformen, die sich auf „social networking“ spezialisiert haben wie Pilze aus dem Boden und tun dies noch heute. Dienste wie „studiVZ“ [stu], die social bookmarking-site „del.icio.us“ [del] oder auch die Onlineenzyklopädie „Wikipedia“

[wik] feierten mehr als beachtliche Erfolge und ein Abebben dieser Welle ist derzeit nicht zu erkennen. Diese Plattformen bieten ein reichhaltiges Betätigungsfeld für die Forscher der oben genannten Fachrichtungen und diese bedienen sich ebenfalls der Werkzeuge, die diese großen Datenmengen in eine übersichtliche und leicht verständliche Form bringen. Nicht zu vergessen sei an dieser Stelle auch die Geographie, die bei der Analyse von geographischen Daten ebenfalls sehr eng mit der Informatik zusammenarbeitet.

1.1 Motivation

Eine der meistgenutzten Visualisierungsmöglichkeiten ist die Darstellung von Daten mittels Graphen, die aus Knoten und Kanten (Bögen) aufgebaut sind. Hierbei stellen Knoten Entitäten von Objekten aus der betrachteten Domäne dar und durch Kanten bzw. Bögen zwischen den Knoten werden Beziehungen zwischen den Objekten verdeutlicht. Um eine solche Visualisierung am Computer umsetzen zu können, wurden in den letzten Jahrzehnten eine Vielzahl von Algorithmen entwickelt, die verschiedene Ansätze und Modelle nutzen, um eine möglichst übersichtliche und verständliche, ja „erhellende“ Darstellung von verschiedensten Daten zu verwirklichen. Der Erfolg dieser automatischen Layout-Algorithmen ist schwer zu messen. Typischerweise wird hier die Effizienz gemessen, mit der ein Computer unter Zuhilfenahme der Algorithmen eine Datenquelle in ein Graphlayout umwandelt. Ein anderer Ansatz ist die Analyse der resultierenden Graphen unter der Berücksichtigung von festgelegten ästhetischen Kriterien. Hier sind zum Beispiel folgende Ästhetikkriterien zu nennen:

- Minimierung der Kantenkreuzungen – Kanten sollen sich nach Möglichkeit nicht kreuzen und sollen immer als eindeutig dieselbe Kante erkannt werden können, auch sollen sie niemals vor oder hinter einem Knoten verlaufen.
- Minimierung der Anzahl der Biegungen/Knicke einer Kante – Alle Kanten sollen eine Mindestlänge besitzen, aber zugleich nicht zu lang sein. Die Zahl der Biegungen bzw. Knicke einer Kante sollte minimal werden, um die Übersichtlichkeit zu wahren.
- Orthogonalität – Der kleinste Winkel bei einer Kantenkreuzung soll maximal werden.

- Kompaktheit der Darstellung – Der Gesamttraum, den der Graph nach seiner Erstellung einnimmt, soll so kompakt wie möglich sein, um eine Übersichtlichkeit zu gewährleisten.

(Dies ist nur eine unvollständige Auflistung der gängigen Kriterien, eine weiterführende Übersicht findet der interessierte Leser in [Pur97], [Pur02] oder in [War02])

Teilweise wurde hier bereits Forschungsarbeit geleistet, wenn es um die Verifizierung dieser Ästhetikkriterien für die Zugänglichkeit und Nützlichkeit für das menschliche Verstehen geht. Sind die nach diesen Kriterien optimierten Graphen tatsächlich in der Lage, die eingebettete Information so darzustellen, dass sie für den menschlichen Betrachter zugänglich ist? Rechtfertigt ein messbarer Erfolg den höheren Aufwand, den eine Optimierung eines automatischen Layoutverfahrens nach diesen Kriterien mit sich bringt? Diese Arbeit wird versuchen, auf diese Fragen Lösungen zu finden, indem hier nicht eine reine Überprüfung der Umsetzung der Ästhetikkriterien erfolgt, sondern experimentell überprüft wird, inwieweit ein automatisch erzeugter Graph für den Betrachter leicht zugänglich ist.

Alternativ zu den Ästhetikkriterien lassen sich auch verschiedene Layoutverfahren miteinander vergleichen. Hier gibt es ähnliche Probleme in der Beurteilung eines „guten“ oder „schlechten“ Layouts. Was macht die Qualität eines Layoutverfahrens aus? Wo liegen die Stärken eines Layouts, wo die Schwächen, unterstützt eine Darstellungsweise einen syntaktischen oder semantischen Zusammenhang besonders gut oder extrem schlecht? Auch hier werden mittels eines experimentellen Verfahrens Antworten gesucht und auch gefunden. Diese Arbeit spezialisiert sich auf drei der gängigsten Layoutverfahren für Knoten-Kanten-Diagramme, nämlich

- ein kräftebasiertes Layout (siehe Abb. 1.1),
- ein orthogonales Layout (siehe Abb. 1.2),
- ein hierarchisches Layout (siehe Abb. 1.3).

(Eine genauere Erläuterung dieser Verfahren ist im theoretischen Teil dieser Arbeit zu finden.)

Ziel dieser Arbeit ist es, die Vor- und Nachteile der verschiedenen Layoutverfahren in Hinsicht auf die Zugänglichkeit für den menschlichen Betrachter offenzulegen. Hierbei kommen verschieden komplexe Graphen zur Verwendung, die jeweils in

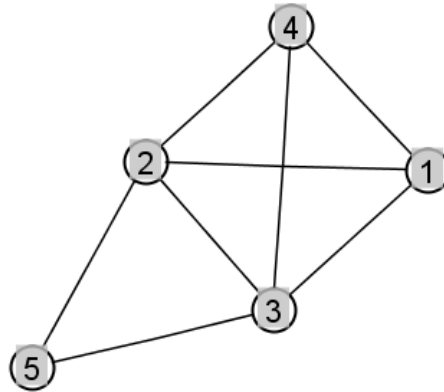


Abbildung 1.1: kräftebasiertes Layout nach Fruchterman und Reingold [Fru91]

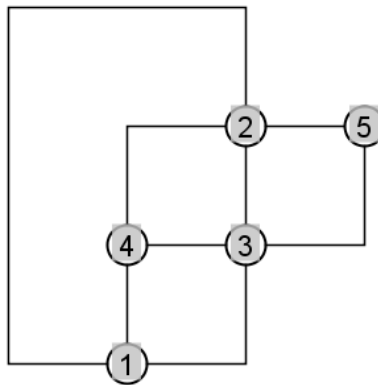


Abbildung 1.2: orthogonales Layout nach Föb Meyer und Kaufmann [Föb95]

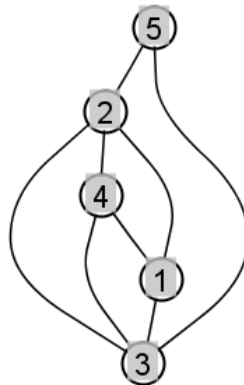


Abbildung 1.3: hierarchisches Layout nach Sugiyama [Sug81]

allen drei Layouts automatisch designed werden und anschließend als Grundlage eines Task-basierten Experiments dienen. Allerdings geht diese Arbeit über eine rein quantitative Analyse der Ergebnisse hinaus, die in bisherigen wissenschaftlichen Abhandlungen zu diesem Thema das Mittel der Wahl war. Hier wird nicht nur eine Analyse der reinen Antwortzeit und der Korrektheit der Antworten stattfinden, sondern darüberhinaus eine qualitative Analyse angestrebt. Diese qualitative Analyse wird durch ein Eyetracking- (Blickregistrierungs-) System realisiert. Dieses System besteht aus einem Monitor, an dem eine Präsentation abläuft, sowie aus einer Kamera, die die Augenbewegungen einer Versuchsperson aufzeichnet und somit festhält, wie die Versuchsperson die gestellten Aufgaben anhand der verschiedenen Graphen abarbeitet.

Denn wenn zum Beispiel bestimmte Kantenkreuzungen essentieller Teil einer Graphdarstellung sind, dann wäre es kontraproduktiv, diese Kreuzung im Zuge eines Optimierungsverfahrens aufzulösen, da so eventuell die Semantik des Graphen verloren gehen könnte. Wie geht man also mit einer solchen (unvermeidlichen) Kantenkreuzung um, wie kann man diese dennoch optimieren? Welches Layout bietet hier die zugänglichste Lösung? Um eine Lösung auf diese Fragen zu finden, muss man verstehen, wie Menschen Graphen lesen.

Um dieses Ziel zu erreichen wurde folglich eine Untersuchung entworfen, die die quantitativen Ergebnisse bisheriger Untersuchungen zu verifizieren versucht, gleichzeitig aber feststellen möchte, wie verschiedene Layouts und deren unterschiedlich ausgeprägte Ästhetikkriterien die „Abarbeitung“ der Aufgaben durch eine menschliche Versuchsperson fördern oder behindern. Das Experiment wurde Task-basiert entworfen, d.h. die Probanden hatten fünf verschiedene Fragen aus dem Bereich der Graphentheorie zu beantworten. Drei verschieden komplexe Graphen wurden jeweils durch die drei oben genannten Layoutverfahren automatisch gestaltet (diese Aufgabe übernahm das Programm DGD, welches am Lehrstuhl für Softwaretechnik an der Universität Trier zu Verfügung steht und dort weiterentwickelt wird), so dass jeder Proband insgesamt 45 Fragen beantworten musste. Nähere Informationen zum Versuchsaufbau und zu DGD sind in Kapitel 2 und 3 zu finden.

Die Resultate dieser Arbeit könnten bei der künftigen Entwicklung neuer Algorithmen bzw. bei der Optimierung bestehender Algorithmen dazu beitragen, dass nicht bedingungslos nach axiomatisch gegebenen ästhetischen Gesichtspunkten optimiert wird, sondern die Zugänglichkeit und Verständlichkeit für einen menschlichen Betrachter im Vordergrund steht.

1.2 Übersicht

Diese Arbeit gliedert sich in zwei Hauptbestandteile. In einem theoretischen Teil werden die benötigten Grundlagen erläutert, die zum Verständnis der späteren Ergebnisse im praktischen Teil vorausgesetzt werden.

In Kapitel 2 werden zunächst die grundlegenden Begriffe der Graphentheorie definiert, wie sie in dieser Arbeit verwendet werden. Im Anschluss daran wird die konkrete Umsetzung dieser Grundlagen in die verwendeten Layouts besprochen. Für alle verwendeten Layoutverfahren (kräftebasiert, orthogonal und hierarchisch) wird eine Definition geliefert und die ästhetischen Ziele erläutert, die hinter den Verfahren stehen. Im Zusammenhang mit diesen Layoutverfahren wird auch auf das verwendete Programm DGD (Dynamic Graph Drawing) eingegangen, welches vom Lehrstuhl für Softwaretechnik der Universität Trier bereitgestellt wird und das das automatische Layouten der Graphen übernommen hat. In einem weiteren Unterkapitel wird das verwendete Eyetracking-System beschrieben und generelle theoretische Hintergrundinformationen zum Thema Eyetracking gegeben. Verwendet wurde hier die Cornea-Reflex-Methode zur Bestimmung der Blickposition auf dem Bildschirm. Den Abschluss des theoretischen Teils wird eine Beschreibung der entworfenen Tasks bilden, die von den Probanden im praktischen Teil zu lösen waren. Hier werden einerseits die graphentheoretischen Grundlagen der Probleme beschrieben und andererseits die Gründe für die Auswahl genau dieser Probleme geliefert.

Kapitel 3 wird sich mit der Auswertung des praktischen Teils dieser Diplomarbeit beschäftigen. Zunächst wird der generelle Versuchsaufbau beschrieben und die Rahmenbedingungen (Stärke und Zusammensetzung der Versuchgruppe, Einweisung in das System und in die Graphentheorie) verdeutlicht. Dann folgt ein kurzer Rückblick auf die Vorstudie zu dieser Versuchsreihe, die ein halbes Jahr vor der eigentlichen Versuchsreihe als eine Machbarkeitsstudie durchgeführt wurde. Die hier mit einer kleinen Personengruppe durchgeführten Tests lieferten wichtige Hinweise zur Versuchspraxis und zu Verbesserungen, die an den Layouts und den Aufgaben vorgenommen wurden. Im Folgenden wird dann eine quantitative Analyse der Ergebnisse des Experimentes durchgeführt, also eine Beurteilung der durchschnittlichen Bearbeitungszeit und der Korrektheit der Antworten aufgeschlüsselt nach Tasks und Layouts. Dies entspricht der Analyse, wie sie bei bisherigen Veröffentlichungen üblich war. Einen Schritt weiter geht die sich anschließende qualitative Analyse des Experiments: mit Hilfe der Daten, die das Eyetracking-System während der Versuche aufgezeichnet hat, ist es möglich, über einen rein quantitativen

Horizont hinaus, die Stärken und Schwächen der einzelnen Layouts bei den verschiedenen Tasks offenzulegen und so detailliertere Ergebnisse zu erhalten.

Im letzten Kapitel werden die Ergebnisse schließlich nochmals zusammengefasst und ein abschließendes Fazit rundet die Analyse ab. Ferner findet sich hier auch eine Reflexion der Erfahrungen, die der Autor im Rahmen dieser Arbeit machen durfte und es wird ein Ausblick gegeben, wie sich künftige Studien anschließen könnten, um weiter offene Fragen zu beantworten oder die Ergebnisse experimentell auf weitere Layouts oder dynamisch über die Zeit veränderliche Graphen zu übertragen.

2 Theoretische Grundlagen

In diesem Kapitel sollen in möglichst kompakter Form die theoretischen Grundlagen der späteren Ergebnisse der experimentellen Untersuchung aufgeführt werden. Hierbei ordnen sich die Unterkapitel bereits in einer Form, die von der allgemeinen Theorie ausgehend, zu immer spezifischeren Hintergrundinformationen im Hinblick auf die Versuchsreihe verläuft. Beginnend mit allgemeiner Graphentheorie wird so übergegangen zur speziellen Graphentheorie, die die verwendeten Layoutalgorithmen beschreibt. Im Hinblick auf die eigentliche Untersuchung werden im Anschluss die theoretischen und technischen Grundlagen des verwendeten Eyetracking-Systems vorgestellt. Sehr eng an der Praxis sind schließlich bereits die Grundlagen der im Experiment verwendeten Tasks, die hier erläutert werden.

2.1 Graphentheoretische Grundlagen

Die Begriffe und grundlegenden Definitionen der Graphentheorie sind glücklicherweise leicht zu verstehen und sollen hier nach Möglichkeit recht kurz abgehandelt werden. Dabei lehnt sich diese Arbeit bei Definitionen und Schreibweisen an [Die06] an, für eine tiefergehende Einführung in die Graphentheorie seien daher eben dieses oder ein beliebiges anderes graphentheoretisches Buch empfohlen. Die Grundlagen und Schreibweisen der Mengentheorie werden als bekannt vorausgesetzt. Daher beginnen wir direkt mit

Definition 2.1 (Graph): Sei V eine endliche Menge. Ein *Graph* G ist ein Paar $G = (V, E)$ disjunkter Mengen mit $E \subseteq [V]^2$, d.h. die Elemente von E sind zweielementige Teilmengen der Menge V . Die Elemente von E heißen dann *Kanten* (bzw. *Bögen*) des Graphen G , während die Elemente von V die *Knoten* des Graphen darstellen.

Diese Definition ist sehr allgemein gehalten und entspricht den in dieser Arbeit betrachteten ungerichteten Graphen. Die Definition eines gerichteten Graphen ent-

spricht dieser Definition mit dem Unterschied, dass die Elemente von E nicht zweielementige Mengen sind, sondern geordnete Paare, so dass die betreffende Kante eine Richtung erhält. Allerdings wird diese Art von Graphen keine Rolle in dieser Arbeit spielen.

Bildlich gesprochen bestehen die hier verwendeten Graphen aus Knoten, die durch Kreise symbolisiert werden, die wiederum zur genauen Identifizierung mit Labels (Großbuchstaben) gekennzeichnet werden. Diese werden durch ungerichtete Kanten verbunden, die sowohl aus geraden Linien (kräftebasiertes Layout), Linienzügen (orthogonales Layout) als auch durch Kurven bzw. Bögen (hierarchisches Layout) dargestellt sind. Diese Unterschiede sind rein ästhetische Variationen und beeinflussen die formale Definition eines Graphen nicht.

Zur Schreibweise bleibt noch anzumerken, dass zur Vereinfachung folgende Konvention getroffen wird: Ein Graph $G = (V, E)$ ist ein Graph auf V , seine Knotenmenge V wird mit $V(G)$ bezeichnet, seine Kantenmenge E mit $E(G)$ und statt $v \in V(G)$ bzw. $e \in E(G)$ schreibt man vereinfacht $v \in G$ bzw. $e \in G$. Je nachdem, ob die Menge $V(G)$ endlich oder unendlich ist, ist auch der Graph G endlich oder unendlich. In dieser Arbeit werden ausschließlich endliche Graphen betrachtet. Die maximal betrachtete Knotenmenge eines hier vorkommenden Graphen wird 20 Knoten betragen. Aus diesem Grund können bei den weiteren Definitionen die Sonderfälle eines unendlichen Graphen oder auch eines leeren Graphen ignoriert werden.

Definition 2.2 (inzident): Ein Knoten $v \in V$ und eine Kante $e \in V \times V$ *inzidieren* miteinander und heißen somit inzident, wenn gilt: $v \in e$. Die beiden Knoten der Kante sind Endknoten und die Kante verbindet diese Knoten.

Für eine Kante (x, y) findet sich auch die einfachere Schreibweise xy bzw. im ungerichteten Fall natürlich auch yx . Ferner heißen zwei Knoten *adjazent* oder benachbart in G , wenn xy oder $yx \in E(G)$ ist. Auch Kanten sind genau dann benachbart, wenn sie einen gemeinsamen Endknoten besitzen. Ein im Rahmen dieser Arbeit sehr wichtiger Begriff im Zusammenhang mit Graphen ist der Begriff der Isomorphie, da gerade die isomorphe Eigenschaft der Versuchsgraphen Rückschlüsse auf die Eignung der verschiedenen Layouts zur Darstellung verschiedener Problemfelder ermöglicht. Aus diesem Grunde folgt nun

Definition 2.3 (Isomorphismus): Sei $G' = (V', E')$ ein weiterer Graph. G heißt genau dann *isomorph* zu G' , in Zeichen $G \simeq G'$, wenn eine Bijektion $\varphi : V \rightarrow V'$ existiert mit $xy \in E \Leftrightarrow \varphi(x)\varphi(y)' \in E'$ für alle $x, y \in V$. Dann ist eine solche Abbildung φ ein *Isomorphismus*.

Der Spezialfall des *Automorphismus* tritt dann ein, wenn gilt: $G = G'$. Um einen möglichen Lerneffekt bei der Versuchsreihe zu minimieren wurde allerdings in dieser Arbeit auf Automorphismen verzichtet, es liegen also ausschließlich echt isomorphe Beziehungen von Graphen gleicher Komplexität (allerdings verschiedener Layouts) vor. In diesem Zusammenhang sei noch der Begriff *abstrakt* erläutert, denn von einem abstrakten Graphen spricht man, wenn es bei der Beschreibung des Graphen nur um seinen Isomorphietyp geht. Eine unter Isomorphie abgeschlossene Klasse von Graphen nennt man ferner eine *Eigenschaft* von Graphen. Im Experiment werden folglich drei Eigenschaften von Graphen zu finden sein, die sich lediglich in ihrer Komplexität unterscheiden.

Ein Task der Versuchsreihe wird sich der Frage widmen, ob ein vorher gezeigter Graph ein Teilgraph eines komplexeren Graphen ist. Um hier ebenfalls eindeutige Begrifflichkeiten zu gewährleisten folgt daher

Definition 2.4 (Teilgraph): Seien $G \cup G' := (V \cup V', E \cup E')$ und $G \cap G' := (V \cap V', E \cap E')$. Ist $G \cap G' = \emptyset$, dann heißen G und G' *disjunkt*. Gilt hingegen $V' \subseteq V$ und $E' \subseteq E$, dann heißt G' ein *Teilgraph* von G bzw. G ist ein *Obergraph* von G' . Man schreibt in diesem Falle $G' \subseteq G$.

Eine weitere Grapheigenschaft, die im Rahmen der zu bearbeitenden Tasks eine wichtige Rolle spielt, ist der Grad eines Knotens:

Definition 2.5 (Grad eines Knotens): Sei $G = (V, E)$ ein Graph. Der *Grad* (alternativ: *Valenz*) eines Knoten v ist die Anzahl $|E(v)|$ der mit v inzidenten Kanten, d.h. da eine Kante immer zwei Endknoten haben muss, ist der Grad eines Knotens die Anzahl seiner Nachbarknoten bzw. die Anzahl der Kanten, deren Endknoten v ist. Man schreibt $d_G(v)$ bzw. $d(v)$.

Eine weitere Eigenschaft von Graphen, die bei der Beschreibung im weiteren Verlauf dieser Arbeit auftreten wird, ist der eng mit der vorgehenden Definition zusammenhängende Begriff des *Durchschnittsgrades*. Der Durchschnittsgrad eines Graphen ist die Zahl

$$d(G) := \sum_{v \in V} d(v) / |V|$$

Diese Zahl misst global, was lokal durch die Knotengrade ausgedrückt werden kann, nämlich eine Annäherung der Anzahl der Kanten von G pro Ecke. Eine direktere Möglichkeit, dieses Verhältnis auszudrücken erhält man durch die Berechnung

$\epsilon(G) := |E|/|V|$. Nun gilt allerdings noch zu beachten, dass bei dieser Zählweise sowohl die Kante xy als auch die Kante yx mitgezählt werden, da die betrachteten Graphen allerdings ungerichtet sind, zählen wir jede Kante doppelt. Daher gilt korrekterweise $\epsilon(G) = \frac{1}{2}d(G)$.

Bei der Beschreibung der auszuführenden Tasks wird auch noch ein weiterer graphentheoretischer Begriff von entscheidender Wichtigkeit sein, nämlich der Begriff des *Weges*. Um sicherzustellen, dass auch bei der Beschreibung dieses Tasks keine Verständnisschwierigkeiten oder Zweideutigkeiten auftreten, sei Folgendes definiert:

Definition 2.6 (Weg): Ein *Weg* (alternativ: *Pfad*) ist ein nicht leerer Graph $P = (V, E)$ der Form

$$V = \{x_0, x_1, \dots, x_k\} \quad E = \{x_0x_1, x_1x_2, \dots, x_{k-1}x_k\},$$

wobei die x_i paarweise verschieden sind. Die Knoten x_0 und x_k sind die Endknoten von P , wobei im Kontext dieser Arbeit x_0 als Startknoten und x_k als Zielknoten bezeichnet werden; sie sind durch P verbunden.

Die Knoten x_2, \dots, x_{k-1} werden als *innere Knoten* des Weges bezeichnet. Wichtig für den angesprochenen Task ist ferner die *Länge* eines Weges, sie ist durch die Anzahl seiner Kanten festgelegt. Einen Weg der Länge k bezeichnet man auch mit P^k . Zur Bezeichnung eines Weges wurden die Probanden dazu angehalten, den Weg durch die natürliche Folge der beteiligten Knoten anzugeben, also in der Form $P = x_0x_1\dots x_k$.

Schließlich existiert auch ein Task, bei dem die Versuchspersonen den Graphen auf die Existenz einer *Clique* überprüfen müssen. Die formale Beschreibung einer Clique folgt daher in

Definition 2.7 (Clique): Sei $G = (V, E)$ ein (ungerichteter) Graph und V' sei eine Teilmenge von V . Dann bezeichnet man V' als *Clique* von G , wenn für jeweils zwei verschiedene Knoten x und y aus V' gilt, dass sie durch eine Kante miteinander verbunden sind.

Das dazugehörige Entscheidungsproblem sieht dermaßen aus, dass es bei einem gegebenen Graphen G und einer natürlichen Zahl k zu entscheiden gilt, ob G eine Clique der Größe mindestens k enthält. Man spricht dann vom *Cliquenproblem*. Dies stellt ein sehr bekanntes *NP*-schweres Problem aus der Komplexitätstheorie

dar (vgl. etwa [Weg07]) und wird in genau dieser Form Bestandteil eines Tasks dieser Versuchreihe sein.

Da im Kontext der ästhetischen Kriterien zur Beurteilung von Graphlayouts der Begriff der *Kantenkreuzung* besonders wichtig ist, sei an dieser Stelle explizit hervorgehoben, dass der in der Graphtheorie bekannte Begriff der Kreuzungsfreiheit in diesem Zusammenhang missverständlich, ja gar fehl am Platze ist. Zwei (oder mehr) Graphen gelten in der Graphtheorie als *kreuzungsfrei*, wenn keiner einen inneren Knoten eines anderen enthält. Im Rahmen dieser Arbeit bezieht sich der Begriff der Kantenkreuzung jedoch nicht auf innere Knoten, sondern auf eine tatsächliche Kreuzung zweier (oder mehrerer) *Kanten* in der Ebene der Abbildung. Wenn also im weiteren Verlauf dieser Arbeit der Begriff *kreuzungsfrei* zu finden sein wird, dann immer im gerade beschriebenen Sinn.

2.2 Graphlayouts

In den folgenden drei Unterkapiteln werden die verwendeten Layoutverfahren dargestellt. Dabei spielen neben einer allgemeinen Beschreibung der Eigenschaften der Layouts insbesondere die algorithmischen Hintergründe eine Rolle, da jedem Layout eine eigene Philosophie zugrunde liegt, die sich bei der konkreten Anwendung in verschiedenen Formen widerspiegelt. Wie im Vorkapitel bereits angesprochen, haben alle Graphen gemeinsam, dass sie aus Knoten und Kanten bestehen, die die Knoten miteinander verbinden.

2.2.1 Das kräftebasierte Layoutverfahren

Das erste Layoutverfahren, das an dieser Stelle vorgestellt werden soll, nimmt so gleich eine gewisse Sonderstellung im Kreise der Layoutalgorithmen ein. Denn während sich andere Layouts in der Regel durch strukturelle Charakteristiken beschreiben lassen, so entspricht das zugrundeliegende Konzept des kräftebasierten Layouts eher einem physikalischen Modell. Dies wird im Folgenden genauer erläutert.

Ein Problem beim Designen eines Graphen besteht darin, einen initialen Zustand zu finden, der dann je nach Verfahren optimiert werden kann. Beim kräftebasierten Layouten ist von Vorteil, dass zu Beginn keinerlei Informationen irgendwie gearteter

struktureller Charakteristika von Nöten ist. Die Anfangsdarstellung des Graphen ist also beliebig bzw. zufällig. Alle nun angesprochenen kräftebasierten Methoden vergleichen den Graphen mit einem System interagierender physikalischer Objekte, deren Ziel es ist, in einen Energie-minimalen Gleichgewichtszustand zu gelangen. Eine weitere große Stärke dieser Ansichtswiese eines Graphen ist, dass die daraus resultierenden Ergebnisse, so die Theorie, in der Regel recht leicht zugänglich sind (was genauer in Kapitel 3 zu beweisen wäre), da dieses Modell unseren tag-täglichen Erfahrungen in der uns umgebenden physikalischen Welt entspricht. Ein weiterer Vorteil ist, dass ein physikalisches Modell recht einfach am Computer programmierbar ist. Außerdem liefert das kräftebasierte Layoutverfahren, insbesondere in kleinen und mittleren Graphen, die wir hier betrachten werden, subjektiv zufriedenstellende Ergebnisse, was die Umsetzung von Ästhetikkriterien angeht: Es findet eine gleichmäßige Verteilung der Knoten in der Ebene statt, die Kanten sind meist von ähnlicher Länge und Knoten, die in Relation stehen (also benachbart sind), werden auch räumlich nahe beieinander abgebildet, jedoch wiederum nicht zu dicht aneinander.

Bei der gleichmäßigen Verteilung von Knoten und dem Streben nach möglichst uniformen Kantenlängen kommt es zu einem Trade-Off: eine möglichst gleichmäßige Verteilung reduziert das Knotenchaos in der Ebene. Allerdings sind dann in der Regel keine annähernd uniformen Kantenlängen zu erreichen, wo es hingegen bei einer Vereinheitlichung der Kantenlängen zur Verzerrung des Graphen kommt. Um dieses Problem zu lösen, brachte Eades den folgenden Vorschlag hervor (vgl. [Ead84]):

„The basic idea is the follows. To embed [lay out] a graph we replace the vertices by steel rings and replace each edge with a spring to form a mechanical system [...] The vertices are placed in some initial layout and let go so that the spring forces on the rings move the system to a minimal energy state.“

Hier modellierte Eades also den Graph als ein physikalisches System aus Stahlingen und Federn, allerdings beachtete sein Modell nicht das Hookesche Gesetz¹. Er passte seinen Ansatz derart an, dass er abstoßende Kräfte zwischen allen Knotenpaaren eines Graphen annahm, allerdings anziehende Kräfte eben nur bei adjazenten Knoten. Der Vorteil dieser Betrachtungsweise ist eine Reduzierung der Zeitkomplexität der Berechnung, da das Berechnen von sich anziehenden Kräften zweier Nachbarknoten in $\Theta(|E|)$ realisiert werden kann, obwohl die Berechnung abstoßender Kräfte $\Theta(|V|^2)$ benötigt. (Für ausführlichere Erklärungen siehe [Fru91] (in dieser Arbeit verwendet) oder [Gre87])

Dieses Modell wurde unter dem Namen „spring embedder“ bekannt, zur Verdeutlichung dient Abbildung 2.1. Die vom Modell benutzten abstoßenden Kräfte zwischen nicht benachbarten Knoten $x, y \in V$ werden charakterisiert durch

$$f_{abs}(p_x, p_y) = \frac{c_\rho}{\|p_y - p_x\|^2} \cdot \vec{p_x p_y},$$

wobei c_ρ eine Abstoßungskonstante ist. Dem gegenüber stehen die Federkräfte

$$f_{fed}(p_x, p_y) = c_\rho \cdot \log \frac{\|p_x - p_y\|}{l} \cdot \vec{p_y p_x},$$

deren Richtung davon abhängt, ob die aktuelle Ausdehnung der Feder größer oder kleiner ihrer natürlichen Länge l ist.

Wie erhält man nun einen Gleichgewichtszustand des Systems? Da die Knotenpositionen anfänglich noch nicht den Positionen entsprechen, die im Gleichgewichtszustand vorgesehen sind, stehen sie bildlich unter Spannung. Um dieses System unter

1 Das Hookesche Gesetz beschreibt das elastische Verhalten von Festkörpern, deren elastische Verformung proportional zur einwirkenden Belastung ist:
Wenn eine Feder zusammengedrückt oder auseinander gezogen und dann losgelassen wird, so wird sie zu ihrer ursprünglichen, oder natürlichen, Länge zurückkehren, vorausgesetzt die Verzerrung war nicht zu groß. Wir sehen, dass für ein kleines Δx die Kraft der Feder annähernd proportional zu Δx_0 ist. Dieses Hookesche Gesetz kann auch folgendermaßen beschrieben werden:

$$F_x = -k(x - x_0) = -kx,$$

wobei k eine Konstante über die Stärke der Feder ist.

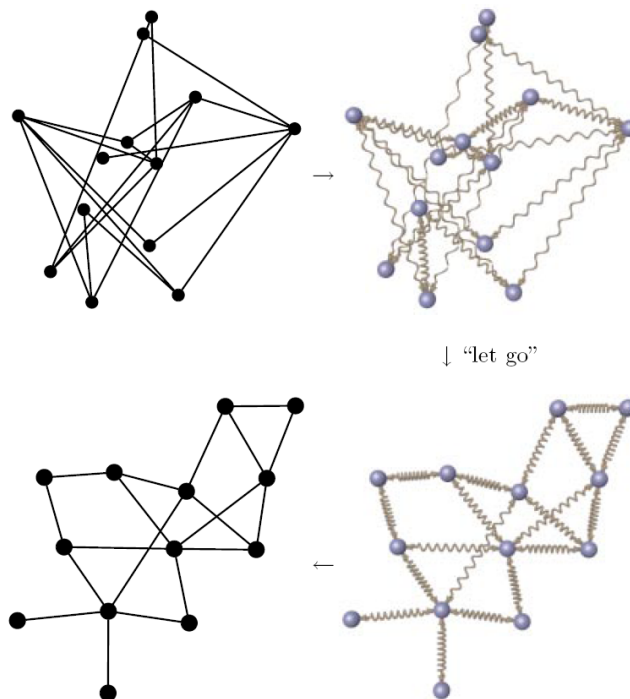


Abbildung 2.1: spring embedder Modell [Bra99]

Spannung in einen Gleichgewichtszustand zu überführen, sieht der entsprechende Algorithmus (vgl. [Bra99]) eine iterative Verschiebung der Knoten zur Zeit t vor, entsprechend eines Kraftvektors $F_y(t)$, der die Summe aller Feder- und abstoßenden Kräfte angibt, die auf den Knoten y einwirken. Nachdem $F_y(t)$ für alle $y \in V$ durchgeführt wurde, wird jeder Knoten eine konstante Anzahl δ mal um diesen Vektor bewegt. Diese Konstante ist von Nöten, um während eines Schrittes ein exzessives Bewegen aufgrund von gleichzeitigen Änderungen zu vermeiden. Sind alle Änderungen abgeschlossen, so befindet sich das System in einem stabilen Zustand, in dem keine lokalen Verbesserungen mehr möglich sind.

Der hier verwandte Algorithmus von Eades ist wie gesehen recht einfach aufgebaut, die Ergebnisse des „spring embedders“ sind allerdings bereits sehr zufriedenstellend im Hinblick auf eine gleichmäßige Verteilung von Knotenmenge und Kantenlängen. Trotzdem setzen auf diesen Grundalgorithmus eine Vielzahl von Verbesserungen auf, deren, aus Sicht des Autors, wichtigste Entwicklungen an dieser Stelle noch in ihren Grundideen erwähnt werden sollen.

Eine Variation von Eades Algorithmus lieferten im Jahre 1988/1989 *Kamada und Kawai* ([Kam88], [Mat89]). Ähnlich wie Eades modellierten auch sie einen Graphen als ein mechanisches System von Federn. Ihr System unterscheidet sich allerdings in dem Punkt vom Eades'schen Modell, an dem letztgenanntes das Hookesche Gesetz ignoriert. An dieser Stelle nutzten Kamada und Kawai partiell differenzierte Gleichungen, um das Layout zu optimieren. Während Eades vorrangig darauf Wert legte, dass zwei adjazente Knoten im Graphen auch in unmittelbarer Nähe abgebildet werden und so, wie bereits erwähnt, anziehende Kräfte nur zwischen benachbarten Knoten betrachtete, ging der Ansatz von Kamada und Kawai einen Schritt weiter: Sie führten das Konzept einer *idealen Distanz* zweier nicht direkt benachbarter Knoten ein. Dabei entspricht die ideale Distanz proportional der Länge des kürzesten Weges zwischen den beiden Knoten. Kamada und Kawai versuchten das Graphproblem optimal zu lösen, indem sie die Gesamtenergie des Systems minimieren, denn wenn die Summe aller Federausdehnungen und -stauchungen minimal wird, dann müssen die Knoten sehr nahe an ihrer idealen Position sein.

Ein weiterer Ansatz zur Verbesserung des Algorithmus von Eades wurde durch Fruchtermann und Reingold in [Fru91] veröffentlicht. Sie betrachten nur zwei Prämissen bei ihren Abbildungen:

1. Knoten, die adjazent sind, sollen nahe beieinander abgebildet werden.
2. Knoten sollen nicht *zu* nahe beieinander abgebildet werden.

Die Nähe der einzelnen Knoten zueinander hängt dabei davon ab, wie viele Knoten abgebildet werden müssen und wieviel Platz dafür zur Verfügung steht. Um hierfür ein vernünftiges Maß zu erhalten, bedienten sich Fruchtermann und Reingold wie Eades zuvor wieder der Natur: Die Kräfte, die zwischen den Knoten wirken verhalten sich genau so wie die Kräfte zwischen Atomen oder die Kräfte zwischen Himmelskörpern. Der entwickelte Algorithmus ähnelt aufgrund dieser Tatsache sehr stark den Algorithmen, die bei molekularen oder astronomischen Simulationen Verwendung finden. Allerdings muss ein entscheidender Unterschied zu diesen Formeln bei Fruchtermann und Reingold beachtet werden: Während im molekularen oder im astronomischen Bereich eine wichtige Kraft die Beschleunigung ist, so gehen die beiden Forscher bei ihrem Modell von linearen Geschwindigkeiten aus. Dieser Unterschied ist zwingend, denn während die naturwissenschaftlichen Formeln zu dynamischen Gleichgewichten (Planetenbahnen etc.) führen, ist im Bereich des Graphzeichnens ein statisches Gleichgewicht erwünscht. Der Algorithmus liefert ebenfalls den vorgegebenen ästhetischen Kriterien in hohem Maße entsprechende Ergebnisse und hat zudem den Vorteil, dass die algorithmische Umsetzung eine Beschleunigung zum Ursprungsalgorithmus darstellt.

Die letzte hier vorgestellte Verfeinerung des ursprünglichen „spring embedder“ - Algorithmus stammt von Frick und wurde in [Fri94] veröffentlicht. Auch hier wurden abstoßende und anziehende Kräfte modifiziert, um so eine Beschleunigung des Algorithmus und eine gleichzeitige Verbesserung der Layoutqualität zu erreichen. Mathematisch gesehen werden die berechnenden Formeln für abstoßende und anziehende Kräfte ohne Quadratwurzel dargestellt und es wird ausschließlich ganzzahlige Arithmetik verwandt. Aus „physikalischer“ Sicht wird eine neue Kraft im System eingefügt, die Gravitation. Sie wirkt auf alle Knoten und zieht diese zu einem Schwerpunkt aller Knoten. Die Stärke dieser Kraft richtet sich nach der Anzahl der Knoten im System. Der Zweck dieser Gravitationskraft ist zu verhindern, dass die einzelnen Bestandteile eines nicht vollständig verbundenen Graphen nicht zu weit auseinander gestoßen werden. Mittels zusätzlicher Heuristiken, die die Ablenkungen oder Rotationen der Knotenbewegungen während der verschiedenen Iterationen des Algorithmus erkennen und verstärken bzw. abschwächen, gelingt es bei diesem Ansatz, die Zahl der Iterationen, die nötig sind, um einen stabilen Zustand zu erreichen, entscheidend zu verringern.

All diese Verbesserungen und Modifikationen des ursprünglichen Algorithmus von Eades zeigen, dass der „spring embedder“ leicht und effizient verändert werden kann, ohne dass der dahinterstehende Grundgedanke des physikalischen Modells verloren geht. Die Ergebnisse der Algorithmen können ausnahmslos als den ästhetischen Kriterien entsprechend angesehen werden, die dem kräftebasierten Layout zugrunde liegen: eine gleichmäßige Verteilung der Knoten, gleichmäßige Kantenlängen und daraus resultierend auch Symmetrie. Ob die Optimierung in Hinsicht auf diese Kriterien allerdings auch die Zugänglichkeit für den menschlichen Betrachter zur Folge hat, wird in Kapitel 3 untersucht werden.

2.2.2 Das orthogonale Layoutverfahren

Das orthogonale Layoutverfahren verfolgt einen völlig anderen Ansatz als das gerade vorgestellte kräftebasierte Verfahren. Im Hinblick auf die Lesbarkeit eines Graphen kann das Vorhandensein von Kanten, die sehr eng beieinander liegen, hinderlich sein. Natürlich stellt es ein Problem dar in direkter Umgebung eines Knotens, an dem sich verschiedene Kanten treffen, die einzelnen Kanten auf Distanz zu halten. Daraus ergibt sich sofort die Betrachtung der Beziehung zwischen den Winkeln zwischen angrenzenden Kanten und der Übersichtlichkeit des Graphen. Genau an dieser Stelle setzt das orthogonale Verfahren an, denn es hat zum Ziel den minimalen Winkel zweier benachbarter Kanten zu maximieren.

Diese Optimierung des Winkels zweier Kanten läuft dabei intuitiverweise darauf hinaus, dass alle Winkel Vielfache von $\frac{\pi}{2}$ sein sollen. Dies hat zur Folge, dass die Kanten achsen-parallel verlaufen. Allerdings hat dieses Verfahren auch den Nachteil, dass dadurch die Knoten nicht mehr länger durch direkte Linien verbunden werden können, sondern die Linien erhalten Knicke bzw. wir betrachten Linienzüge. Daher ist es ein weiteres Ziel eines „guten“ orthogonalen Layouts, die Anzahl der Kantenknicke zu minimieren. Diese und weitere Vorgaben führten zu einer sehr produktiven Forschung im Bereich des orthogonalen Graphen-Zeichnens, aus der eine Vielzahl von Methoden und Heuristiken hervorgingen, deren Grundlagen nun in diesem Kapitel erläutert werden sollen.

Wie bereits angesprochen ist das Grundproblem, das hier im Vordergrund steht, wenn es um eine zugängliche Darstellung eines komplexen Graphen geht, das folgende:

Problem 2.1 (Winkelauflösung): Sei $G = (V, E)$ gegeben mit n Knoten und m Kanten. Wie kann man dann die Knoten als Punkte einer Ebene und die Kanten als gerade Verbindungslinien zwischen adjazenten Knoten anordnen, so dass der kleinste Winkel zwischen zwei benachbarten Knoten maximal wird?

Zu dieser Problemstellung wurden im Laufe der Jahre eine Vielzahl von unterschiedlichen Lösungsansätzen publiziert. Von Formann wurde in [For90] gezeigt, dass das Problem zu prüfen, ob ein planarer Graph mit höchstem Grad 4 so gezeichnet werden kann, so dass der jeweils kleinste Winkel mindestens $\frac{\pi}{2}$ beträgt, NP -hart ist. Weitere Untersuchungen im Bezug auf die Lösung des Problems mit direkten Kanten wurden auch von Malitz und Papakostas in [Mal94] durchgeführt, wobei immer genauere Schranken hinsichtlich der Komplexität des Problems herausgefunden wurden. Ein ausführlicherer Überblick würde den Rahmen dieser Arbeit sprengen, kann jedoch in [Eig99] nachvollzogen werden.

Allerdings entwickelte sich auch eine alternative Sichtweise für dieses Problem, die die einfache Unterscheidbarkeit der einzelnen Kanten des Graphen als wichtiges Kriterium ansah: Wenn man es erlaubt, die Kanten nicht mehr als direkte Linien zu zeichnen, sondern als Pfade, die aus mehreren Kantensegmenten bestehen, so ist es möglich alle Kanten achsenparallel auszurichten und wir erhalten eine orthogonale Darstellung. Dafür müssen wir allerdings in Kauf nehmen, dass der Abbildung zusätzliche (virtuelle) Knoten hinzugefügt werden, immer an der Stelle, wo ein Pfad seine Richtung ändert. Wir definieren daher

Definition 2.8 (orthogonales Raster): Eine *Einbettung* Γ eines Graphen $G = (V, E)$ in ein orthogonales Raster ist eine Abbildung in die Ebene, die Knoten $x \in V$ auf ganzzahlige Gitterpunkte $\Gamma(x)$ und Kanten $(x, y) \in E$ auf Pfade im Gitter abbildet, so dass die Endpunkte $\Gamma(x)$ und $\Gamma(y)$ miteinander verbunden sind. Eine Gittereinbettung heißt einfach, wenn die Anzahl der Knicke null ist. Eine einfache Einbettung induziert eine mögliche Partition der Kantenmenge E in eine horizontale Menge E_h und eine vertikale Menge E_v .

Die Ausrichtung der Kanten in einer orthogonalen Zeichnung ist dabei wohl - strukturiert, da es nur zwei Klassen von Kantensegmenten gibt und die Segmente jeder einzelnen Kante wechseln zwischen der horizontalen und der vertikalen Klasse. Dadurch werden für eine solche Abbildung sehr positive visuelle und strukturelle Eigenschaften möglich und eine enge Verwandtschaft zum Bereich des VLSI Layout Designs wird deutlich, wobei die Komponenten eines Micro-Chips den Knoten und deren Verbindungen den Kanten des Graphen entsprechen. Es gibt ferner noch eine Gemeinsamkeit der beiden Disziplinen: Bei beiden sind Kanten- bzw. Verbindungskreuzungen hochgradig unerwünscht, so dass die Vermeidung von Kantenkreuzungen für beide Gebiete zu einer der wichtigsten Prämissen wurde.

Bisher haben wir allerdings nur ein grobes Modell eines orthogonalen Layouts. Um einen konkreten Graphen orthogonal zu zeichnen, brauchen wir mehr: Wir müssen die Knicke entlang einer Kante beschreiben können, inklusive der Richtung, die die Kante am betreffenden Knick einschlägt. Hierzu sind einige Vorüberlegungen nötig: Ein Graph $G = (V, E)$ kann als Liste von Knoten und als verbindende Kanten angesehen werden. In der Ebene allerdings kommen zusätzliche topologische Informationen dazu. Durch die Kanten des Graphen wird die Ebene in eine Reihe von Gebieten (*faces*) unterteilt. Die Struktur dieser Gebiete wird durch die sie umgebenden Kantenzüge charakterisiert. Man kann erkennen, dass Gebiete und deren Nachbargebiete durch die zirkuläre Anordnung der Kanten eines jeden Knotens bestimmt werden. Daher können wir nun die Einbettung eines Graphen in die Ebene beschreiben durch eine Liste von Knoten V , eine Liste von Kanten E , eine Liste von Gebieten (*faces*) F und eine Liste $P(f)$ von Kanten für jedes Gebiet f .

Zurück zum eigentlichen Problem des Beschreibens einer Kante mitsamt ihrer Knicke: Für jede Kante in der Kantenliste $P(f)$ beschreiben wir die Reihenfolge der Knicke durch einen binären String, während wir dem Verlauf der Kante folgen. Hierbei wird ein Linksknick als „1“ und ein Rechtsknick als „0“ kodiert. Hat eine Kante keine Knicke, so wird ihr der leere String „ε“ zugeordnet. Schließlich ordnet man jeder Kante einen Winkel (ein Vielfaches von $\frac{\pi}{2}$) zu, der begrenzt wird durch ihr letztes Kantensegment und das erste Segment der nächsten Kante. Diese Art der

Kodierung nennt man *orthogonale Repräsentation* eines Graphen. Die orthogonale Repräsentation eines Graphen enthält nur kombinatorische und einige geometrische Informationen über den Graphen, wohingegen keine Aussagen über die Länge von Kanten integriert sind. Aufgrund der Wohl-Definiertheit der orthogonalen Repräsentation für jeden Graphen, kann man einige notwendige Bedingungen erkennen:

- Es existiert ein Graph entsprechend der Listen V und E .
- Jede Kante ist doppelt kodiert, für jedes Gebiet, das sie begrenzt, einmal. Diese beiden Kodierungen müssen dabei konsistent sein.
- Die Summe aller Winkel eines Gebietes f_i wird beschrieben als $H(f_i)$ und muss im Hinblick auf die Tatsache, dass f_i ein geradliniges Polygon ist, konsistent sein.
- Für jeden Knoten x muss sich die Summe der Winkel von aufeinanderfolgenden Ecken um x zu 4 aufsummieren.

Nun haben wir die Werkzeuge an der Hand, um die Grundversion des zugrundeliegenden Algorithmus nachvollziehen zu können. Wir wollen feststellen, wie wir jedem Kantensegment, das in die orthogonale Repräsentation eingebunden ist, eine ganzzahlige Länge zuordnen können. Diese Aufgabe wurde von Di Battista in [Bat99] als eine dreiphasige *topology-shape-metrics* Methode präsentiert. Die erste Phase dieser Methode hat zum Ziel, eine Topologie der Abbildung durch das Ermitteln einer kombinatorischen Einbettung und der Außenflächen des Graphen festzulegen. In der zweiten Phase wird die Form der orthogonalen Zeichnung fixiert, wobei das Ergebnis dieser Phase eine orthogonale Repräsentation H ist. Die dritte Phase des Algorithmus legt die Metriken der Abbildung fest, so dass daraus resultierend eine Einbettung in das orthogonale Raster entsteht. Diese letzte Phase des Algorithmus soll jetzt noch etwas genauer dargelegt werden.

Die Idee hinter dieser Methode (vgl. [Tam87]) ist, wie bereits im einleitenden Text erwähnt, das Hinzufügen von künstlichen Knoten zu H , so dass auf diese Weise leichter eine Abbildung Γ' der resultierenden orthogonalen Abbildung H' gefunden werden kann. Hierbei gilt es zu beachten, dass wir uns beim Hinzufügen künstlicher Knoten immer noch auf der Ebene der orthogonalen Repräsentation befinden und bisher keine geometrischen Aktionen stattfinden. Durch das Entfernen der künstlichen Knoten gelangen wir erst zu einer Einbettung Γ in das orthogonale Raster für die ursprüngliche Repräsentation H . Wir betrachten noch genauer die Einzelheiten

dieser Methode: Durch das Ersetzen jedes Knickes (die Anzahl der Knicke sei b) in H durch einen künstliche Knoten hat unsere orthogonale Repräsentation $n + b$ Knoten. Dann wird für jedes daraus resultierende Gebiet f in H der binäre String $S(f)$ generiert, wie oben bereits angedeutet: Die Kanten eines Gebietes werden im Uhrzeigersinn betrachtet und abhängig vom Winkel $a(e)$ (der ein Vielfaches von $\frac{\pi}{2}$ ist) wird dem String eine „0“ (wenn $a(e) = 1$), eine „1“ (wenn $a(e) = 3$) oder „11“ (wenn $a(e) = 4$) hinzugefügt, $a(e) = 2$ kann ignoriert werden, da dann kein Knick der Kante vorliegt. Dieser String repräsentiert schließlich das jeweilige Gebiet f . Um „Umwege“ zu vermeiden, werden die Strings noch so optimiert, dass alle Vorkommen von „100“ durch „0“ ersetzt werden. Anhand des Strings kann auch festgestellt werden, ob das Gebiet f ein inneres Gebiet (in diesem Fall ist $S(f) = „0000“$) oder äußeres Gebiet (dann enthält der String keine zwei aufeinander folgenden Nullen) ist.

Man findet nun eine orthogonale Einbettung von H' durch Zuweisen gleicher Längen für gegenüberliegende Seiten der rechteckigen Gebiete. Hier kommt ein Flussalgorithmus zur Anwendung. Die Idee hinter diesem Algorithmus ist die Berechnung zweier minimaler Kostenflüsse, wobei ein Fluss die vertikalen Kanten, der andere die horizontalen Kanten abdeckt. Für jede Richtung wird also ein Netzwerk (N_h und N_v) gebildet, so dass die Vereinigung der Netzwerke der duale Graph von H' ist. Jeder Bogen eines Netzwerkes entspricht einer e in H' und der Fluss durch a wird wiedergegeben durch die Länge von e . Die untere Schranke für den Fluss durch a ist 1 und wir verwenden uniforme Flusskosten. Die Flussershaltung garantiert nun, dass gegenüberliegende Seiten die gleiche Länge haben müssen. Die Flusskosten addieren sich zur Gesamtlänge der Kante und durch die Minimierung der Kosten wird das Ergebnis optimal. Zur Verdeutlichung dieses Gedankens betrachten wir Abbildung 2.2. Durch diese Methode wird eine optimale Ausdehnung des Graphen gewährleistet, allerdings können keine Aussagen über maximale Kantenlängen gemacht werden.

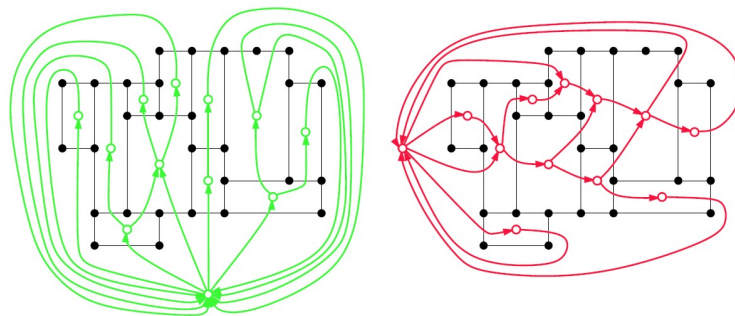


Abbildung 2.2: Die Netzwerke N_v und N_h [Eig99]

Wie beim kräftebasierten Layout, so gibt es auch hier eine Vielzahl von Heuristiken und Verfeinerungen, die den Algorithmus weiter optimieren sollen. Eine ausführliche Beschreibung dieser Erweiterungen würde den Rahmen dieser Arbeit sprengen, daher seien hier nur einige wenige mit ihren Grundideen genannt. Der 4-phasige Algorithmus von Tamassia und Tollis [Bat89] konstruiert ebenfalls eine Einbettung in das orthogonale Gitter für einen Graph G . Dieser Algorithmus hat eine lineare Laufzeit und die resultierende Graphzeichnung minimiert die Anzahl der gesamten Kantenknicke des Graphen auf maximal $2,4 \cdot n + 2$ Knicke, wobei keine Kante mehr als 4 Knicke hat. Auch über die Länge der Kanten lässt der Algorithmus eine Aussage zu, diese ist $O(n)$ und der Gesamttraum, den die Einbettung annehmen kann, ist $O(n^2)$. Der Algorithmus von Biedl und Kant [Bie94] verfolgt folgende Vorgehensweise: Zunächst wird eine st -Ordnung (vgl. Definition 6.4 in [Eig99]) des Eingabegraphen berechnet, um dann die Knoten nacheinander entsprechend dieser Reihenfolge einzeln in das Gitter einzuordnen. Dabei wird für jeden Knoten eine neue Reihe zum Layout hinzugefügt und für jede noch nicht vollständige Kante eine neue Spalte. Durch diese Technik kann eine Gittergröße von höchstens $(m - n + 1) \cdot n$ gewährleistet werden.

Über die Jahre entstanden so eine große Anzahl von Methoden, um eine orthogonale Zeichnung möglichst optimal im Hinblick auf wünschenswerte Eigenschaften wie die Minimierung von Kreuzungen oder der Anzahl von Knicken zu gestalten. Ob diese Optimierungen tatsächlich die Zugänglichkeit für einen menschlichen Betrachter einer orthogonalen Einbettung eines Graphen unterstützen, soll zumindest im Rahmen der hier betrachteten Tasks in Kapitel 3 beleuchtet werden.

2.2.3 Das hierarchische Layoutverfahren

Bereits Ende der 1970er Jahre entstanden die ersten Ideen einer hierarchischen Art des Graphenzeichnens. Die ersten Ansätze dieses Layoutverfahrens findet man bei Warfield [War77] und auch kurze Zeit später bei Carpano in [Car80]. Die bekannteste Methode des hierarchischen Graphenzeichnens wurde dann im Jahre 1981 von Sugiyama veröffentlicht [Sug81], diese wurde 1990 von Eades und Sugiyama [Ead90] nochmals erweitert. Ähnlich wie es beim orthogonalen Layoutverfahren der Fall war und ist, so wurde auch bei dieser Methode des Graphenzeichnens sehr viel Forschungs- und Entwicklungsarbeit in den letzten Jahren geleistet, die in einer großen Anzahl von Heuristiken, Methoden und Implementierungen mündete und dies noch heute tut. In fast allen Werkzeugen zur automatischen Gestaltung von Graphen ist das hierarchische Layout als wichtige Option vertreten, so auch

beim Programm DGD, das in dieser Arbeit die Aufgabe des Layoutens der Graphen übernimmt.

Beim hierarchischen Zeichnen von Graphen spielt die Richtung einer Kante eine große Rolle, vielleicht wesentlich stärker als es bei den vorausgegangenen Layoutverfahren der Fall war, denn ein Qualitätskriterium eines hierarchischen Layouts ist die Eigenschaft, dass Kanten nach Möglichkeit immer von oben nach unten zeigen sollen, umgekehrte Kanten sind unerwünscht. In dieser Arbeit werden allerdings ausschließlich ungerichtete Graphen verwendet, daher gelten einige Einschränkungen, die im späteren Verlauf dieses Kapitels auftreten werden, hier nur bedingt. Um allerdings die Grundstruktur des hierarchischen Layoutens verstehen zu können, werden diese Bedingungen hier zumindest erwähnt.

Neben dem angesprochenen Ästhetikkriterium gibt es aber noch einige weitere, daher soll hier zunächst eine Übersicht gegeben werden:

- Kanten, die nach oben zeigen, sollen vermieden werden.
- Die Knoten sollen gleichmäßig verteilt sein.
- Lange Kanten sollen vermieden werden.
- Die Zahl der Kantenkreuzungen soll minimiert werden.
- Kanten sollen möglichst geradlinig und vertikal verlaufen.

Wie bei den bereits abgehandelten Layoutalgorithmen festgestellt, ist eine gleichzeitige Optimierung für alle diese Designziele unmöglich, da diese sich teilweise gegenseitig ausschließen. Aber bereits der Versuch auch nur einige dieser Kriterien zu erfüllen, erweist sich als *NP*-hart. Daher geht man beim Layoutvorgang den Weg des Teilens des Verfahrens in vier Schritte. Die erste Phase des Algorithmus stellt das *Entfernen von Zyklen* im Graph dar. Hierzu werden bei so wenigen Kanten wie möglich die Richtung geändert bis der Graph azyklisch ist. Dadurch wird es möglich die Kanten alle in eine Richtung verlaufend zu zeichnen, was in der nächsten Phase wichtig sein wird. Allerdings sei an dieser Stelle nochmals daran erinnert, dass bei der hier vorliegenden Arbeit ausschließlich ungerichtete Graphen betrachtet werden, daher ist dieser Schritt im vorliegenden Fall nicht von Nöten. Die zweite Phase befasst sich mit der *Berechnung einer Hierarchie*. Um hier eine klare Begrifflichkeit zu haben, folgt an dieser Stelle zunächst

Definition 2.9 (*n*-stufige Hierarchie): Sei $G = (V, E)$ ein Graph. Eine Partition L der Knotenmenge V von G in L_1, \dots, L_h heißt *n*-stufige Hierarchie. Die Mengen L_1, \dots, L_h heißen dann die *Schichten* der Hierarchie.

Eine Hierarchie ist also eine Unterteilung der Knotenmenge eines Graphen in disjunkte Mengen. Dabei entsprechen alle Knoten einer Partition einer Schicht im Layout des Graphen. Wir werden im weiteren Verlauf für viele Algorithmen Hierarchien benötigen, die eine besondere Eigenschaft aufweisen, diese werden *einfache* Hierarchien genannt. Eine Hierarchie ist genau dann *einfach*, wenn nur Kanten zwischen benachbarten Schichten auftreten. Um dieses zu gewährleisten, führt man das Prizip der *Dummy* - Knoten ein, falls adjazente Knoten mehr als eine Schicht auseinander liegen.

Die dritte Phase des Verfahrens ist die sehr wichtige Phase der Reduzierung der Kantenkreuzungen durch *Anordnung der Knoten*. Hierzu werden im weiteren Verlauf einige gängige Ansätze vorgestellt werden. Die Grundidee ist, dass für jede Schicht eine Ordnung der darin befindlichen Knoten berechnet wird, wobei diese Berechnung das Ziel hat, eine möglichst kreuzungsminimale Ordnung zu entwerfen. In der Regel geschieht dies durch ein Überprüfen von benachbarten Schichten und der dazwischen verlaufenden Kanten. Dazu später mehr.

Die vierte Phase des hierarchischen Layoutverfahrens hat zur Aufgabe, die *Knotenpositionen* zu *berechnen* und anschließend die Kanten zu platzieren. Die Kriterien, die es hier zu beachten gilt, sind, dass sich die Knoten nicht gegenseitig überlagern und dass nach Möglichkeit kein Knoten so positioniert wird, dass er auf dem geraden Weg zwischen zwei adjazenten Knoten liegt. Hat jeder Knoten seine Position gefunden, so werden abschließend die Kanten, im hier vorliegenden Fall als Kurven, eingezeichnet.

Bei der genaueren Beschreibung der vier Phasen wird aus bekannten Gründen auf eine Erklärung der ersten Phase verzichtet. Bei Interesse der hier zum Tragen kommenden Heuristiken ist eine gute Übersicht bei [Bas99] zu finden. Wir befassen uns an dieser Stelle mit der zweiten Phase. Sei L eine Partition von V , z.B. $L = L_1, \dots, L_h, \bigcup_{i=1}^h L_i = V$. Die *charakteristische Funktion* dieser Partition wird bezeichnet durch $y : V \rightarrow 1, \dots, h$, d.h. $y(u) = i$ wenn $u \in L_i$. In diesem Kontext ist auch der Begriff der *Spanne* einzuführen: Die Spanne der Kante (u, v) ist gegeben durch $y(u) - y(v)$. Nun kann die Vorgehensweise zur Schichteinteilung, wie bei [San96] geschehen, formuliert werden:

- Berechne die charakteristische Funktion y durch Tiefensuche oder Breitensuche, wodurch eine beliebige Partitionierung in Linearzeit entsteht.
- Für den Fall, dass die Kanten gewichtet sind, muss an dieser Stelle noch eine Minimalkostenoptimierung stattfinden, dies ist jedoch bei den hier betrachteten Graphen nicht der Fall.
- Benutze einen „spring embedder“ (vgl. Kapitel 2.2.1). Hierbei müssen die bereits vorgestellten f_{fed} und f_{abs} beachtet werden. Allerdings ist nur die vertikale y -Koordinate bei der Berechnung von Interesse.

Erweiterungen dieses Algorithmus führen zu einer Minimierung der Höhe der Partitionierung, selbst unter Beachtung einer maximalen Breite. Siehe hierzu den *Coffman-Graham-Algorithmus* in [Cof72]. Weitere Optimierungen haben die Erfüllung des Ästhetikkriteriums „minimiere die Spannen der Kanten“ zum Ziel. Dieses Ziel kann gleichgesetzt werden mit der Minimierung der Anzahl der dummy - Knoten, was nach [Ead90] auch eine minimale Hierarchiehöhe garantiert: Die Eigenschaften einer Hierarchie können folgendermaßen angegeben werden.

$$\begin{aligned} y(u) - y(v) &\geq 1 \quad \forall (u,v) \in E \\ y(v) &\in \mathbb{Z}^+ \quad \forall v \in V \end{aligned}$$

Eine Minimierung über $\sum_{(u,v) \in E} (y(u) - y(v))$ minimiert die Gesamtspanne der Kanten und damit auch die Zahl der dummy - Knoten.

Die dritte Phase des hierarchischen Layoutverfahrens ist die Anordnung der Knoten und die damit verbundene Minimierung der Kantenkreuzungen. Hierbei gilt es festzustellen, dass die Anzahl der Kantenkreuzungen nicht von der exakten Position der Knoten abhängt (diese wird erst in der nächsten Phase festgelegt), sondern nur von der relativen Position eines Knotens im Verhältnis zu den anderen Knoten, also der Knotenordnung. Obwohl wir also nicht die exakten x -Koordinaten betrachten müssen, erweist sich das Problem dennoch als *NP*-hart (vgl. [Bas99]). Bei der Vielzahl der Methoden, die entwickelt wurden, um die Zahl der Kantenkreuzungen zu minimieren, gibt es nur wenige, die das Problem global minimieren, die meisten optimieren lokal die Kantenkreuzungen zwischen jeweils zwei Schichten (*Layer - by - Layer - Sweep*). Beim *Layer - by - Layer - Sweep* wird zunächst eine Knotenordnung mittel BFS oder DFS bestimmt, anschließend wird eine Schicht gewählt, deren Knotenordnung unberührt bleibt, während die Knoten der nächsten Schicht

so angeordnet werden, dass sich die Zahl der Kantenkreuzungen zwischen beiden Schichten minimiert. So durchläuft die Methode alle Schichten nacheinander.

Auch hier verfolgen verschiedene Heuristiken diverse Ansätze. Die bekannte *Barycenter* Heuristik [Sug81] geht z.B. davon aus, dass insbesondere in Graphen mit wenigen Kanten jeder Knoten nahe bei seinen adjazenten Knoten plaziert werden sollte. Die *Median* Heuristik hingegen bestimmt die x -Koordinate jedes Knoten durch Berechnung der Median - Werte seiner Nachbarknoten (vgl. [Ead94]). Zur Übersicht über die Bubblesort - ähnliche *Greedy Switch* Heuristik, die qualitativ hochwertige aber laufzeitintensive *Split* Heuristik, die *Optimal Crossing Minimization* und ähnliche Methoden empfiehlt sich wiederum die Lektüre von [Bas99].

Am Ende des hierarchischen Layoutverfahrens steht die exakte Bestimmung der x -Koordinaten der einzelnen Knoten in Phase vier. Ziel hier ist es, die Anzahl der Biegungen möglichst klein zu halten. Wir haben bereits gesehen, dass diese Zahl von der Anzahl der dummy - Knoten abhängt. Ein Ansatz zur Lösung dieses Problems liefert [Gan93], indem er einen Hilfsgraph einführt, durch den das Problem der x -Positionierung in das bereits gelöste Problem des Anordnens der Knoten in einer Schicht transformiert wird. Dadurch, dass er den Graphen dreht und die x -Positionierung wie eine Schicht behandelt. Die Positionierung der Kanten, die in dieser Phase ebenfalls ansteht, stellt im Kontext der vorliegenden Arbeit ein leicht zu lösendes Problem dar, da alle Knotendarstellungen sowohl eine uniforme Größe als auch Form haben. Für den Fall unterschiedlicher Formen bzw. Größen sei im Hinblick auf dieses Problem einmal mehr auf [Bas99] verwiesen.

Ob all diese Optimierungen und ausgeklügelten Heuristiken tatsächlich dazu beitragen, einen Graph leichter zugänglich zu machen und wo die spezifischen Stärken und Schwächen des hierarchischen Layoutverfahrens zu finden sind, wird Kapitel 3 zeigen.

2.2.4 DGD – Dynamic Graph Drawing

Nachdem die in dieser Arbeit verwendeten Layoutverfahren nun so knapp wie möglich, aber so ausführlich wie nötig, vorgestellt wurden, folgt jetzt eine Beschreibung des Tools, das für die Umsetzung dieser Algorithmen verantwortlich zeichnet. Dieses trägt den Namen DGD (Dynamic Graph Drawing) und wird vom Lehrstuhl für Softwaretechnik des Fachbereichs Informatik der Universität Trier zur Verfügung gestellt und weiterentwickelt. Dieses Programm bietet eine Implementierung

aller drei angesprochenen Layoutverfahren und lieferte somit für die Aufgabe des automatischen Layoutens der Graphen dieser Arbeit eine ideale Plattform. Die Berechnung eines Layouts von vorgegebenen Graphen, wie es für diese Arbeit benötigt wird, nutzt jedoch nur einen Bruchteil der Möglichkeiten von DGD aus, dessen eigentliche Aufgabe die Visualisierung dynamischer Graphen im Verlauf der Zeit ist. Hier gibt es eine Vielzahl von Visualisierungsmöglichkeiten und veränderbaren Parametern, sowie Analysemöglichkeiten, die z.B. im Bereich der Analyse dynamischer Netzwerke voll zur Geltung kommen. Trotz dieser Vielfalt von Einsatzmöglichkeiten ist das Programm darauf optimiert, die Darstellung auf einem Standard - Monitor zu ermöglichen und dies mit geringen Hardware - Anforderungen zu realisieren.

Das Hauptfenster von DGD präsentiert zunächst beim Start eine leere Arbeitsfläche. Diese kann dann durch Zeichnen per Hand mit dem gewünschten Graph gefüllt werden. Alternativ kann ein vorgefertigter Graph von einem Medium in DGD geladen werden. Das hier verwendete Format *gaml* ist ein auf XML basierendes Dateiformat, das sich ableitet von Grundelementen von GraphML, einem Dateiformat zur Beschreibung der strukturellen Eigenschaften eines Graphen. Neben der Arbeitsfläche bietet DGD die nötigen Bedienelemente, um durch eine dynamische Darstellung eines Graphen zu navigieren. Für die Gestaltung der hier verwendeten Graphen weit wichtiger ist der Menüpunkt „Optionen“, denn hier kann man nicht nur die Umstellung zwischen gerichteten und ungerichteten Graphen vornehmen, sondern dem Graph auch ein Layout zuweisen, das über Parameter noch feinjustiert werden kann.

Die in DGD implementierten Layoutalgorithmen sind darauf optimiert, neben einer hohen Qualität im Hinblick auf die Erfüllung ästhetischer Kriterien, außerdem eine große Layoutstabilität zu gewährleisten. Dies ist im Kontext dieser Arbeit, die sich ausschließlich mit statischen Graphen beschäftigt, von geringerer Bedeutung, bei der Betrachtung von dynamischen Graphen dient dies jedoch der Erhaltung der „*Mental Map*“ (vgl. [Mis95]). Durch die hochwertige Implementierung der ästhetischen Kriterien der Layoutalgorithmen entstehen jedoch Graphen, die für die Verwendung in dieser Versuchsreihe prädestiniert sind. Wie die Algorithmen konkret implementiert sind, soll im Folgenden kurz angedeutet werden:

Der Layouter für das kräftebasierte Layoutverfahren benutzt die Methode von Fruchterman und Reingold [Fru91], die bereits in Kapitel 2.2.1 angesprochen wurde. Implementiert ist ein „*spring embedder*“ mit Gravitation und *simulated annealing*. Die Knoten des Graphen werden durch anziehenden Kräfte zwischen adjazenten Knoten und abstoßenden Kräften zwischen allen Objekten positioniert. Eine Erweiterung des Ansatzes von Brandes [Bra02] stellt der orthogonale Layouter dar.

Dieser wiederum basiert auf dem Netzwerk - basierten Ansatz von Fößmeier und Kaufmann [Föß95]. Wie in Kapitel 2.2.2 beschrieben, wird der Graph in ein orthogonales Raster abgebildet. Der Layouter des hierarchischen Layouts basiert auf dem Ansatz von Sugiyama [Sug81], wie er in Kapitel 2.2.3 vorgestellt wurde. Zunächst wird eine Verteilung der Knotenpartition gleichmäßig auf die Schichten der Hierarchie nach den genannten Kriterien verteilt. Dieses Verfahren ordnet jedem beliebigen Graph eine hierarchische Darstellung zu, so dass kein bereits mit einer hierarchischen Ordnung versehener Graph vorausgesetzt wird, was bei dieser Untersuchung optimal ist.

2.3 Eyetracking – Grundlagen und Technik

In diesem Unterkapitel sollen einige Begrifflichkeiten definiert werden, um bei der späteren Betrachtung der Ergebnisse die entsprechenden Termini verwenden zu können. Im Rahmen dieser Arbeit kann nur ein grober Überblick auf diese sehr umfassende Thematik gegeben werden. Für weiterreichende Informationen können die Werke von Duchowsky [Duc02], Leven [Lev91] oder auch von Neumann [Neu99] herangezogen werden.

Es gibt bei der Aufzeichnung von Augenbewegungen eine Vielzahl von Methoden und Techniken, die sich in den betrachteten Bewegungsindikatoren unterscheiden. Zum Beispiel werden bei der Elektro-Okulographie mittels Hautelektroden elektrische Potentialdifferenzen um die Augen abgegriffen, die durch die Augenbewegungen erzeugt werden. Bei der „Mannheimer Blickregistrierung“ wird die Augenpartie der Versuchsperson gefilmt und aus der Stellung des Auges berechnet, welcher Bereich gerade betrachtet wird. Die bei dieser Untersuchung verwendete Methode der Blickregistrierung ist die *Cornea-Reflex-Methode*. Diese Methode basiert auf der lichtreflektierenden Eigenschaft der Hornhaut (Cornea) des Auges und nutzt so die sich bei der Augenbewegung verändernden Ausfallswinkel der Reflektion. Hierbei sind zwei Verfahren der Cornea-Reflex-Methode zu unterscheiden. Bei Benutzung eines head - mounted Systems ist die (Infrarot-) Lichtquelle am Kopf der Versuchsperson befestigt. Hierdurch wird eine größere Bewegungsfreiheit ermöglicht. Da bei der vorliegenden Untersuchung aufgrund des Versuchsaufbaus jedoch keine größeren Kopfbewegungen zu erwarten waren, wurde hier auf die Verwendung einer am Monitor befindlichen Kamera inklusive Lichtquelle gesetzt (remote eyetracking). Die Positionierung ist für die messtheoretischen Grundlagen nicht entscheidend.

Die anatomischen Grundlagen der Cornea-Reflex-Methode beruhen auf der Tatsache, dass die Cornea, im Gegensatz zum kugelförmigen Rest des Augapfels, aus dem Auge herausragt und somit eine elliptische Form hat. Daher hat die Cornea an verschiedenen Punkten verschiedene Krümmungsradien. Wird nun das Auge bewegt, so verschiebt sich die Lage der Cornea und die Reflektion des auf das Auge gerichteten Lichtstrahls wird verändert. Diese Richtungsänderung des reflektierten Lichtstrahls kann nun verwendet werden, um die Größe und Richtung der Augenbewegung zu berechnen, da zwischen Reflexionswinkel und Augenbewegung eine Gesetzmäßigkeit besteht. In Abbildung 2.3 wird ein Beispiel einer Cornea Reflexion gezeigt.

Befasst man sich mit dem Thema Eyetracking, so stößt man sehr schnell auf den Begriff der *Saccaden*. Als Saccaden werden extrem schnelle Augenbewegungen bezeichnet, die den Zweck haben, ein neues Detail eines betrachteten Objektes mittels einer Fixation auf der Sehgrube (Fovea Centralis) des Auges abzubilden. Diese Bewegungen geschehen nicht langsam zum neuen Details hingerrichtet, sondern das Auge springt ruckartig zu diesem Detail und verharrt anschließend dort. Man unterscheidet hier zwischen *automatisch gestarteten* und *willentlich initiierten* Saccaden. Erstere treten im Rahmen von plötzlichen, unwillkürlichen neuronalen Aufmerksamkeitsreaktionen auf. Letztgenannte Saccaden dienen dazu, ein bisher periphär wahrgenommenes Objektdetail genauer zu untersuchen. Da Saccaden extrem schnell sind, stellt sich die Frage, ob diese überhaupt für die Blickbewegungsregistrierung von Belang sind, bzw. ob durch Saccaden überhaupt Informationen aufgenommen werden können. Hierbei erweist sich das Ergebnis mehrere Untersuchungen als interessant, das besagt, dass während einer Saccade die Sensitivität des Sehvermögens heruntergefahren wird, man ist zwar nicht blind, aber die Fähigkeit, Details zu erkennen, geht verloren. Auf der anderen Seite kann gezeigt werden, dass es, ohne Saccaden zwischen den *Fixationen*, zur Überlagerung der Informationen kommt, die während der Fixationen aufgenommen werden. Daher sind die Saccaden ein essentieller Bestandteil der Blickbewegungsregistrierung.



Abbildung 2.3: Beispiel einer Cornea Reflexion

Nun wurde bereits ein weiterer essentieller Begriff genannt, wenn man sich mit Eye-tracking - Studien befasst: *Fixation*. Bei einer Fixation ruht der Blick bewegungslos auf einem Detail des betrachteten Objektes. Über die Dauer, die der Blick verharren muss, um als Fixation zu gelten, gibt es in der Literatur sehr unterschiedliche Angaben zwischen 50 msec. und bis zu 500 msec. Für die genaue Festlegung einer oberen oder unteren Grenze für eine Fixationsdauer gibt es keine Nachweise oder theoretische Begründungen. Man orientiert sich daher an einem Wert von 50 msec., um von einer Fixation zu sprechen. Einheitlich ist allerdings die Meinung, dass es während einer Fixationsphase zur eigentlichen Informationsaufnahme kommt. Das heißt, während einer Fixation kommt es zur Aufnahme *und* Verarbeitung der Information. Damit bilden Fixationen auch im Rahmen dieser Untersuchung das entscheidende Kriterium zur Beurteilung der Ergebnisse in der qualitativen Analyse.

2.4 Die verwendeten Tasks

Als es um die Planung dieser Untersuchung ging, waren eine Vielzahl von Variablen zu klären. Unter anderem musste zunächst einmal festgelegt werden, welchen grundsätzlichen Aufbau die Versuchreihe haben sollte. Hier war z.B. vorstellbar, entweder rein syntaktische oder mit einer Semantik versehene Graphen zu verwenden. Eine weitere Frage war, ob eher ein taskorientierter Ansatz gewählt werden sollte oder ob die Probanden nach ihren Präferenzen zu den einzelnen Layouts befragt werden sollten. Um möglichst aussagekräftige Ergebnisse zu erhalten, wurde im Hinblick auf das Thema dieser Arbeit, nämlich der Vergleich verschiedener Layoutverfahren, ein rein syntaktischer Ansatz gewählt, d.h. die Graphen wurden neutral, ohne eine semantische Einbettung, in den verschiedenen Layouts gezeichnet. Außerdem entschieden wir uns für einen taskbasierten Ansatz, da es mit Hilfe des Eyetracking Systems möglich wird, die Präferenzen der Probanden implizit mit zu erfassen. In diesem Kapitel sollen daher die Tasks und deren theoretische Hintergründe erläutert werden.

2.4.1 Task 1 – Existenz eines Knotens

Task 1 hat ein wenig den Charakter einer „Aufwärm - Frage“. Der genaue Text der Aufgabe lautet:

Untersuchen Sie, ob ein Knoten mit der Bezeichnung „B“ in dem folgenden Graph vorhanden ist.

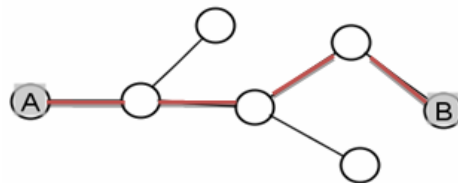
Diese Frage soll von den Probanden mündlich beantwortet werden, indem sie dies bejahen oder verneinen. Die Bezeichnung des Knotens ist von Frage zu Frage verschieden, allerdings wird bei allen Graphen gleicher Komplexität immer der jeweils entsprechende Knoten gesucht. Task 1 kann als \exists -quantifiziert angesehen werden, denn eine Antwort ist bereits dann möglich, wenn der Knoten gefunden ist, es müssen nur im worst case alle Knoten betrachtet werden. Außerdem müssen bei dieser Frage keinerlei Beziehungen (Kanten) zwischen Knoten untersucht werden, diese können gänzlich ausgeblendet werden. Bei dieser Frage werden aufgrund des geringen Schwierigkeitsgrades sehr kurze Antwortzeiten und korrekte Antworten erwartet. Dennoch kann sie sehr interessante Ergebnisse hervorbringen, denn es wird wichtig sein zu sehen, ob eines der Layouts Vorteile gegenüber den anderen hat, wenn es z.B. darum geht, einen Punkt in einer Karte zu finden. Aufgrund der insgesamt recht langen Testphase ist es außerdem erholsam für die Probanden, zwischenzeitlich Fragen weniger hohen Niveaus zu bearbeiten.

2.4.2 Task 2 – Wegeproblem

Bereits eine deutliche Steigerung des Schwierigkeitsgrades liefert Task 2:

Ein Weg von Knoten A nach Knoten B existiert dann, wenn es möglich ist, über (beliebig viele) mit Kanten verbundene Knoten von A aus den Knoten B zu erreichen.

Beispiel:



Finden Sie heraus, ob im nun folgenden Graphen ein Weg von Knoten „I“ nach Knoten „F“ existiert und nennen Sie diesen!

Dieser Task entspricht dem aus der Komplexitätstheorie bekannten REACHABILITY- oder Erreichbarkeitsproblem. Das Erreichbarkeitsproblem ist ein *NL*-vollständiges

Problem. Es lässt sich beispielsweise mit Hilfe der Breitensuche oder der Tiefensuche lösen. Da wir hier ausschließlich ungerichtete Graphen betrachten, könnte man die Frage, ob ein Zielknoten von einem Startknoten aus erreichbar ist auch dadurch lösen, indem man untersucht, ob der Graph zusammenhängend ist.

Die Frage geht allerdings bei genauer Betrachtung über das reine Entscheidungsproblem hinaus, das mit „ja“ oder „nein“ zu beantworten wäre. Zusätzlich wird vom Probanden noch erwartet, dass er den Weg auch nennt, sofern er einen finden kann. Auf diese Weise ist es bei der späteren Auswertung möglich, die Fehlerquellen leichter zu bestimmen, da so der Fehler jeweils auf eine bestimmte Kante eingeschränkt werden kann. Wie bei Task 1 so sind auch hier die Bezeichnungen von Start- und Zielknoten je nach Layout verschieden, allerdings entsprechen sich die Knoten jeweils gleich komplexer Graphen. Eine weitere Gemeinsamkeit zu Task 1 ist, dass diese Frage als \exists -quantifiziert zu bewerten ist, denn auch hier müssen nicht alle Knoten betrachtet werden, die Suche kann eingestellt werden, sobald ein Weg gefunden ist.

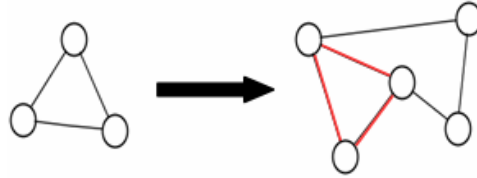
Eine Besonderheit dieser Frage ist die Tatsache, dass nicht nach dem kürzesten Weg zwischen den beiden Knoten gefragt wird, wie es bei ähnlichen Untersuchungen bisher in der Regel der Fall war, sondern der Proband kann ohne Vorgabe einen beliebigen Weg bestimmen. Durch diese Liberalisierung wird es möglich sein festzustellen, ob ein Layout implizit den kürzesten Weg bei uniformen Pfadkosten liefert, also für eine solche Spezialaufgabe prädestiniert wäre. Der Schwierigkeitsgrad dieser Aufgabe wird höher eingeschätzt als bei Task 1. Es wird ferner vermutet, dass sowohl die quantitative als auch die qualitative Analyse deutliche Unterschiede zwischen den Layouts hervorbringen werden, da hier den Kanten als Verbindungslinien eine wichtige Rolle zukommt, welche jedoch die charakteristischsten Unterschiede der drei betrachteten Layoutverfahren ausmachen.

2.4.3 Task 3 – Existenz eines Teilgraphen

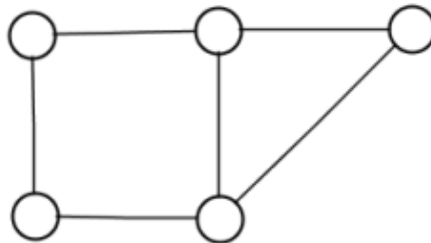
Bei Task 3 haben die Probanden die Aufgabe zu untersuchen, ob der Graph ein Obergraph eines vorher gezeigten Graphen ist:

Ein gleicher Teilgraph liegt dann vor, wenn in zwei Graphen Knoten existieren, die in Anzahl und Verbindungen (Kanten) untereinander gleich sind.

Beispiel:



Untersuchen Sie, ob der Graph



in irgendeiner Form im folgenden Graphen vorhanden ist und nennen Sie die beteiligten Knoten.

Genaugenommen ist ein Teilgraph eines Graphen G ein Graph, der durch Entfernen von beliebigen Knoten und Kanten aus G entsteht (bei Entfernen eines Knotens x fallen auch alle mit x inzidenten Kanten weg). Hier offenbart sich allerdings ein Grundproblem, das bei jeder Aufgabenstellung bestand: Um einer heterogenen Gruppe von Versuchspersonen bei den Fragestellungen gerecht zu werden, wurde zugunsten der Verständlichkeit die graphentheoretische Präzision heruntergesetzt.

Auch dieser Task ist \exists -quantifiziert, denn es reicht, nur so viele Knoten zu betrachten, bis man einen Teilgraph gefunden hat. Es wird also nicht nach allen Teilgraphen (und bei manchen Graphen gibt es tatsächlich mehrere Möglichkeiten) gefragt, sondern die Angabe einer Lösung reicht. Zusätzlich zum Entscheidungsproblem sollen wie bei Task 2 auch die fünf beteiligten Knoten genannt werden.

Der Schwierigkeitsgrad dieser Aufgabe wird als sehr hoch eingeschätzt, da allein der Begriff „Teilgraph“ einem ungeübten Probanden schwer zu vermitteln ist. Den Versuchspersonen muss verdeutlicht werden, dass sie nicht nach einer bestimmten Form, einem „Umriss“, suchen sollen, sondern dass eine rein syntaktische Eigenschaft des

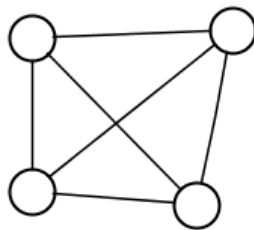
Graphen gefragt ist. Die Form spielt hier also keine Rolle. Für den Probanden ist es also bereits beim Durchlesen der Frage wichtig, die Eigenschaften der Vorlage zu erfassen, sich diese zu merken und erst dann im eigentlichen Graph danach zu suchen. Zur Vereinfachung werden bei allen Graphen, egal welcher Komplexität, der gleiche potentielle Teilgraph gezeigt. Es liegt die Vermutung nahe, dass trotz dieser Erklärungen das kräftebasierte Layout aufgrund seiner geraden Kanten hier einen Vorteil haben wird.

2.4.4 Task 4 – Cliques-Problem

Eng verwandt mit Task 3 ist Task 4:

Eine Clique ist eine Knotenkonstellation, bei der jeder Knoten mit jedem anderen Knoten der Clique durch eine Kante verbunden ist.

Beispiel:



Existiert in folgendem Graphen eine *4er-Clique*? Wenn ja, nennen Sie bitte die beteiligten Knoten.

Dieses Entscheidungsproblem ist ein sehr bekanntes *NP*-vollständiges Problem (vgl. [Weg07]) und rechtfertigt somit seine Aufnahme in diese Studie. Die korrekte Definition des Cliquesproblem lautet:

Sei $G = (V, E)$ ein Graph. Existiert zu G und einer natürlichen Zahl k eine Clique in G der Größe mindestens k . Das heißt, die Frage ist, ob G mindestens k Knoten enthält, die untereinander paarweise verbunden sind. Wie bei beiden vorherigen Tasks reicht die bloße Bejahung oder Verneinung der Frage nicht aus, sondern die beteiligten Knoten der gefundenen Lösung müssen ebenfalls angegeben werden. Dies dient auch hier einer effizienten und zielgerichteten Suche nach den Stärken und Schwächen der einzelnen Layouts. Wie seine Vorgänger kann auch Task 4 als \exists -quantifiziert angesehen werden, denn es genügt eine Clique zu finden und zu benennen.

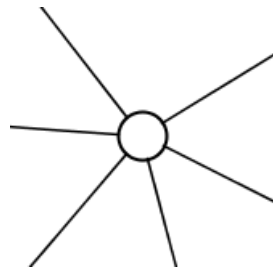
Der Schwierigkeitsgrad wird vergleichbar hoch wie bei Task 3 eingeschätzt, da die beiden Tasks, wie bereits erwähnt, eine enge Verwandtschaft aufweisen. Man könnte sagen, Task 4 ist ein Spezialfall von Task 3, denn eine 4er - Clique kann als spezieller Teilgraph angesehen werden. Die Ergebnisse werden zeigen, ob die beiden Tasks sich dennoch signifikant unterscheiden, denn die Fragestellung zu Frage 4 beschreibt den Teilgraph direkt mit seinen Eigenschaften, während diese bei Frage 3 vom Probanden selbst erkannt werden müssen.

2.4.5 Task 5 – Maximaler Grad

Den Abschluss des Task - Kanons bildet Task 5:

Der Grad eines Knotens ist die Anzahl der von diesem Knoten ausgehenden Kanten.

Beispiel (Grad 6):



Finden Sie im folgenden Graphen den Knoten mit dem *höchsten Grad*, nennen Sie diesen und geben Sie die Anzahl der ausgehenden Kanten an!

Dieser Task nimmt im Vergleich zu den anderen vorgestellten Aufgaben eine gewisse Sonderrolle ein, denn um diese Frage zu beantworten, muss man, zumindest in der Theorie, alle Knoten betrachten und kann erst dann eine sicher korrekte Antwort geben. Hier liegt also ein \forall -quantifizierter Charakter einer Frage vor. Hier wird interessant sein, ob die Probanden tatsächlich jeden einzelnen Knoten betrachten oder ob einzelne Knoten bei einem bestimmten Layout keine Beachtung finden. Die Schwierigkeit der Aufgabe wird geringer eingeschätzt als bei Task 2 bis 4, da hier keine Beziehungen zwischen Knoten wichtig sind, sondern zur Beantwortung reicht es, jeden Knoten für sich und dessen unmittelbare Umgebung zu betrachten.

Trotzdem eignet sich dieser Task ebenfalls sehr gut zum Vergleich der Layouqualität, da die verschiedenen Verfahren gerade bei der Darstellung von Kanten in der Nähe eines Knotens große Unterschiede aufweisen (hier sei an die Grundidee des orthogonalen Layouts (vg. Kapitel 2.2.2) erinnert). Welches Konzept hier für den Menschen am zugänglichsten ist, wird im nun folgenden Kapitel 3 behandelt werden.

3 Durchführung und Auswertung der Studie

In diesem Kapitel wird zunächst eine Übersicht darüber gegeben, unter welchen Bedingungen die Versuchsreihe durchgeführt wird. Hierzu gehören der Versuchsaufbau inklusive einer Beschreibung der technischen Ausstattung, Informationen über die Versuchsgruppe sowie über den eigentlichen Ablauf des Experimentes. Bevor dann die quantitative Auswertung der Daten stattfindet, findet ein kurzer Rückblick auf die vorausgegangene Vorstudie statt, die als Machbarkeitsstudie für die jetzige Untersuchung ausgelegt war. Hierbei ergaben sich einige interessante Thesen, die es in der Hauptstudie zu bestätigen oder zu widerlegen gilt, auch beim Versuchsaufbau wurden Schwachstellen erkannt und beseitigt. In der folgenden quantitativen Auswertung werden Erkenntnisse mit Hilfe von statistischen Auswertungen über die Korrektheit der Antworten und die benötigte Bearbeitungszeit gewonnen. Zum Schluss des Kapitels werden die Daten, die vom Eyetracker aufgezeichnet werden, ausgewertet und liefern so eine Grundlage für die qualitative Analyse.

3.1 Versuchsaufbau und -rahmenbedingungen

Um eine solche Studie durchzuführen sind eine Reihe von Vorüberlegungen von Nöten. Welche Voraussetzungen müssen gegeben sein, damit die Ergebnisse der Untersuchung aussagekräftig sind? Wie viele Probanden sollten teilnehmen und wie viel Belastung kann man ihnen zumuten? Welche Zeitvorgaben sollen festgelegt werden? Welche Einschränkungen gibt es auf der technischen Seite zu beachten? Hier wird zunächst ein Überblick über die Zusammensetzung der Versuchsgruppe gegeben:

3.1.1 Probanden

Zur Teilnahme an dieser Evaluation wurden 36 Probanden herangezogen. Es handelt sich hierbei ausschließlich um Studenten der Universität Trier, die sich freiwillig für die Teilnahme an der Studie gemeldet haben. Es wurden im Voraus keine körperlichen Einschränkungen als Ausschlusskriterium festgelegt, da insbesondere das Tragen einer Sehhilfe, sei es eine Brille oder Kontaktlinsen, keine technischen Probleme bei der verwendeten Eyetracker - Hardware verursacht. Dies wurde bei einer Vorstudie bereits überprüft. Allerdings werden alle körperlichen Einschränkungen, die einen Einfluss auf das Ergebnis haben könnten, mittels eines Fragebogens erfasst. Durch diesen Fragebogen lassen sich auch weitere Informationen über die Zusammensetzung der Versuchsgruppe sammeln.

Die Gruppe setzt sich zusammen aus jeweils sechs Politologen und Geographen, jeweils vier Informatikern, Psychologie - Studierenden und Studierenden der Betriebswirtschaftslehre, je drei Juristen und Anglistikstudierenden, zwei Pädagogen, zwei Medienwissenschaftlern und jeweils einem Studierenden der Germanistik und der Japanologie. Diese breite Streuung ist explizit gewünscht und ganz im Sinne des Ziels dieser Studie, denn die hier erhaltenen Ergebnisse sollen allgemeingültig sein. Dies ist nur bei einer möglichst heterogenen Versuchsgruppe zu erreichen. Der Vergleich verschiedener Graphlayouts kann dabei für fast alle diese Fachrichtungen wichtige Impulse liefern, denn egal ob Geograph oder Politologe, ob Betriebswissenschaftler oder Informatiker, jeder wird in seinem Berufsleben tagtäglich mit Graphen konfrontiert, die der Visualisierung von verschiedensten Themen dienen. Eine Layoutverbesserung kann hier gebietsübergreifend zu entscheidenden Verbesserungen führen. Bei einer Zusammensetzung der Gruppe ausschließlich aus Informatikstudenten wäre, aufgrund eines sehr engen Themabezuges und der damit verbundenen einseitigen Sichtweise auf das Thema Graphen, kaum eine allgemeingültige Übertragung der Ergebnisse möglich gewesen. Außerdem kann nur so gewährleistet werden, dass Probanden mit unterschiedlichster Vorbildung an der Studie teilnehmen.

Eine weitere Charakteristik der Versuchsgruppe ist die Verteilung der Geschlechter: An der Studie nahmen 22 Frauen und 14 Männer teil. Der Frauenanteil liegt damit bei 61%. Auch hier ist die wünschenswerte 50:50 - Verteilung fast erreicht, die Abweichung liegt im Akzeptanzrahmen. Beim Alter zeichnet sich die erwartete homogene Gruppe ab. Alle Probanden sind altersmäßig in einer Spanne von 9 Jahren, der Jüngste mit 20 und der Älteste mit 29 Jahren. Das Durchschnittsalter der Versuchspersonen beträgt 23,5 Jahre. Dies ist bei einer Untersuchung, die ausschließlich Studenten als Probanden anspricht zu erwarten und stellt keine starke

Einschränkung der Ergebnisse dar. Eine weitere Streuung ist bei zukünftigen Umfragen wünschenswert. Die Probanden wurden auch befragt, inwiefern „deutsch“ ihre Muttersprache ist. Dies ist besonders wichtig für die Verwertbarkeit der Ergebnisse, da ein Verstehen der gestellten Fragen essentiell für dieses Experiment ist. 95% der Teilnehmer der Studie sind Personen mit „deutsch“ als Muttersprache, bei der kleinen Gruppe mit fremdsprachlichem Hintergrund wird verstärkt darauf geachtet, dass es zu keinerlei Verständnisschwierigkeiten kommt. Die Deutschkenntnisse dieser Gruppe sind jedoch durchweg sehr gut, so dass hier keine Schwierigkeiten zu erwarten waren.

Im Verlauf der Studie traten bei keiner Versuchsperson grundsätzliche Verständnisschwierigkeiten im Bezug auf die Tasks auf. Alle Probanden wirkten sehr engagiert und versuchten mit einer hohen Motivation, die ihnen gestellten Aufgaben zu lösen. Auch im Anschluss an die eigentliche Befragung zeigte ein großer Teil der Studenten ein reges Interesse an den Inhalten und dem Zweck der Studie. Dies demonstriert, dass das Thema dieser Studie durchaus im allgemeinen Interesse liegt. Die Gesamtdauer, die eine Versuchsperson für das gesamte Experiment inklusive Vor- und Nachbereitung benötigte, lag bei ca. 45 Minuten. Jeder Proband erhielt im Anschluss eine Aufwandsentschädigung von 5€.

3.1.2 Design der Untersuchung

Das Design dieser Studie verfolgt als ein primäres Ziel, allgemeingültige, aussagekräftige Ergebnisse für den Vergleich verschiedener Graphlayouts zum automatischen Design von Graphen zu liefern. Hierzu werden zunächst Graphen benötigt, die das sinnvolle Layouten nach den verschiedenen Verfahren ermöglichen, die bestimmte Eigenschaften besitzen, wie sie für die sinnvolle Bearbeitung der Tasks nötig sind und schließlich sollte die Komplexität so gewählt werden, dass die Informationen des Graphen noch sinnvoll von den Probanden zu erfassen sind.

Um diesen Anforderungen gerecht zu werden, wurden drei Graphen „per Hand“ entworfen. Diese unterscheiden sich in ihrer Komplexität: ein Graph besteht aus 10 Knoten und 15 Kanten, der zweite aus 15 Knoten und 23 Kanten und der dritte aus 20 Knoten und 27 Kanten. Der durchschnittliche Grad bei allen Graphen ist also etwa konstant 3. Dieser Wert wurde festgelegt, um einerseits eine gute Übersichtlichkeit zu gewährleisten, auf der anderen Seite war es mit dieser Vorgabe möglich, die gewünschten Eigenschaften (Existenz eines bestimmten Knotens, Existenz eines nicht trivialen Weges, Existenz eines Teilgraphen, Existenz einer Clique, Existenz

einer nicht trivialen Lösung für die Frage nach dem höchsten Knotengrad) zu realisieren. Zur eindeutigen Identifizierung der Knoten wird jedem Knoten ein Namen gegeben, der aus einem Großbuchstaben besteht. Die Kanten werden ungerichtet gezeichnet. Ist ein Graph von Hand gestaltet, so wird er anschließend in DGD übertragen und durch Zuweisung eines Layoutalgorithmus findet ein Layouten des manuellen Graphen entsprechend des Layoutverfahrens statt. Auf diese Weise entstehen aus jedem der drei Grundgraphen jeweils drei unterschiedlich gezeichnete Graphen, die jedoch syntaktisch die gleichen Informationen enthalten. Um einen Lerneffekt oder ein Wiedererkennen in der späteren Studie auszuschließen, werden nun die Knotennamen bei gleich komplexen Graphen permutiert. Auf diese Weise ist es nicht mehr ersichtlich, welche Knoten einander in den verschiedenen Layouts entsprechen. Damit ist ein „Merken“ einer Lösung oder ein „Auswendiglernen“ bestimmter Knoten nutzlos und kann die Ergebnisse nicht verfälschen. Entsprechend der Permutationen der Knotennamen müssen natürlich auch die Fragen zu den verschiedenen Tasks angepasst werden. Hier ist besonders darauf zu achten, dass Task 1 sich bei jedem Graphen gleicher Komplexität auf den entsprechenden Knoten bezieht und dass bei Task 2 jeweils äquivalente Start- und Endknoten angegeben werden. Diese Spezialisierung der Fragen ist bei den Tasks 3 - 5 nicht nötig, da hier in der Fragestellung keine festen Knotennamen vorhanden sind. Eine Übersicht über die verwendeten Graphen findet man im Anhang.

Sind die Graphen wie beschrieben fertig gestaltet und die Namenspermutationen abgeschlossen, dann werden die Graphen den entsprechenden Tasks zugeordnet. Da durch die jeweils drei Layouts der drei Grundgraphen insgesamt neun verschiedene Abbildungen entstanden sind und für jede dieser Abbildungen alle Tasks beantwortet werden müssen, sind insgesamt bei der Evaluation pro Proband 45 Fragen zu beantworten. Im Hinblick auf die Benutzung des Eyetrackingsystems findet eine Gruppierung von jeweils 15 Fragen (und der damit verknüpften Abbildungen) zu einem Frageblock statt. Diese Maßnahme wurde gewählt, um dem Probanden bei der Evaluation nach 15 Fragen die Möglichkeit einzuräumen, sich wieder zu entspannen, um keine Verfälschung der Ergebnisse durch eine zu hohe Dauerbelastung zu riskieren. Die Vorstudie, bei der mit unterschiedlich großen Frageblöcken experimentiert wurde, bestätigt, dass eine Blockgröße von 15 Fragen den optimalen Kompromiss darstellt zwischen einer zu langen Konzentrationsphase und einer zu häufigen Neukalibrierung, was das konzentrierte Arbeiten zu oft unterbrochen hätte. In jedem Frageblock steht die Reihenfolge der Tasks fest: Die Tasks werden in ihrer Reihenfolge von 1 - 5 durchlaufen, nach Beantwortung von Task 5 beginnt die nächste Runde wieder bei Task 1 und so weiter. Insgesamt sind also pro Frageblock drei Fragedurchläufe zu bewältigen. Während die Reihenfolge der Tasks gleich bleibt, werden die zu den Fragen gezeigten Abbildungen permutiert.

Das heißt, dass bei der Festlegung der Reihenfolge der verschiedenen Abbildungen darauf geachtet wird, dass innerhalb einer Fragerunde nach Möglichkeit immer alle drei Layouts und alle drei Graphkomplexitäten auftauchen, wobei auch hier die Reihenfolge permutiert wird, um einen Lerneffekt noch weiter auszuschließen. Die konkrete Reihenfolge der Abbildungen kann Tabelle 3.1 entnommen werden.

Task Nr.	Frageblock 1	Frageblock 2	Frageblock 3
Task 1	10 hier.	15 kräfte.	20 ortho.
Task 2	20 ortho.	10 ortho.	10 kräfte.
Task 3	20 hier.	20 ortho.	10 hier.
Task 4	15 ortho.	20 kräfte.	20 hier.
Task 5	10 kräfte.	10 hier.	15 kräfte.
Task 1	10 ortho.	15 hier.	20 kräfte.
Task 2	15 kräfte.	15 ortho.	20 hier.
Task 3	20 kräfte.	10 kräfte.	15 kräfte.
Task 4	15 hier.	20 ortho.	10 ortho.
Task 5	20 ortho.	20 hier.	15 hier.
Task 1	15 ortho.	20 hier.	10 kräfte.
Task 2	10 hier.	15 hier.	20 kräfte.
Task 3	15 hier.	10 ortho.	15 ortho.
Task 4	10 kräfte.	15 kräfte.	10 hier.
Task 5	20 kräfte.	15 ortho.	10 ortho.

Tabelle 3.1: Reihenfolge der Graphen in den Frageblöcken 1 - 3

Nachdem also die Reihenfolge feststeht, in der die Graphen und die Tasks gezeigt werden, muss dieser Entwurf in die Eyetracking Software übertragen werden. Dabei gilt es die Vorgaben und Erfahrungswerte umzusetzen, die die Vorstudie lieferte. Die in diesem Fall benutzte Software CLEARVIEW 2.5.1 bietet eine Reihe von Präsentationsformen [Tob05]:

1. Bilder - (Image - Stimulus): Bei dieser Präsentationsform wird dem Probanden eine Bilddatei am Monitor gezeigt.
2. Bildfolge - (Slideshow - Stimulus): Diese Form ist vergleichbar mit einer Microsoft Powerpoint Präsentation und dafür gedacht, mehrere Bilder aneinanderzureihen. Diese Bilder werden dem Probanden dann in der vorgegebenen Reihenfolge entweder bei Tastendruck oder zeitgesteuert am Monitor eingeblendet.

3. Internetseiten - (Web - Stimulus): Hier können Webseiten als Testgrundlage verwendet werden. Dazu muss die URL der gewünschten Startseite im Internet, oder einer auf der Festplatte gespeicherten, browserfähigen Datei angegeben werden. Wird die Präsentation gestartet, öffnet sich automatisch der Browser und die Testperson kann alle Browserfunktionen verwenden.
4. Bildschirmaufzeichnung - (Screen - Stimulus): Der Screen - Modus eignet sich zur Präsentation externer Software-Anwendungen. ClearView zeichnet alle Vorgänge auf dem Präsentationsmonitor als Video auf.
5. Bewegte Bilder - (Moving Bitmaps - Stimulus): Diese Präsentationsform wird verwendet, um eine oder mehrere Bilddateien auf dem Bildschirm zu animieren. Dazu wird ein Skript in Form einer Textdatei benötigt, in dem die Bewegungsrichtung vorgegeben wird.
6. Video - (AVI - Stimulus): Der AVI-Modus bietet die Möglichkeit, den Testpersonen kleine Filme zu zeigen.

Für die hier vorliegende Untersuchung eignet sich besonders die Präsentationsform „Slideshow - Stimulus“, da hier eine sehr individuelle Gestaltung eines jeden Frageblocks möglich ist, bei einer gleichzeitigen Gewährleistung, dass für jedes einzelne Vorkommen eines Tasks separate Eyetracking Daten aufgezeichnet werden. Für jeden Frageblock wird ein Slideshow - Stimulus angelegt. In diesen werden die Fragen entsprechend der Reihenfolge in Tabelle 3.1 auf die Art gespeichert, dass die jeweilige Frage erst ausgeblendet wird, wenn der Proband sie gelesen und verstanden hat. Hier gibt es kein Zeitlimit, daher wird erst bei Tastendruck die nächste Folie erscheinen. Diese wiederum ist nicht der Graph, auf den sich die Frage bezieht, sondern zwischen jede Frage und den entsprechenden Graph wird ein Fixationskreuz eingebaut. Dieses ist notwendig, um den Ausgangspunkt der Blickbewegung für alle Probanden gleich zu halten und so vergleichbare Ergebnisse zu erhalten. Diese Folie besteht aus einem kleinen schwarzen Kreuz auf weißem Hintergrund, das zentral auf dem Monitor angeordnet ist. Diese Folie wird automatisch nach einer Dauer von drei Sekunden durch den eigentlichen Graph ersetzt, so dass die Bearbeitung stattfinden kann.

Um einen möglichen Übungseffekt noch weiter zu minimieren, als dies durch die beschriebenen Permutationen bereits geschieht, wird zusätzlich noch die Reihenfolge der Frageblöcke 1 - 3 randomisiert. Es ergeben sich sechs verschiedene Kombinationen, in denen die Frageblöcke bearbeitet werden können. Auf diese Weise kann

sichergestellt werden, dass im Falle unterschiedlicher Schwierigkeitsgrade der Frageblöcke (was im Test nicht nachgewiesen werden konnte) keine Vor- oder Nachteile entstehen, weder für den einzelnen Probanden noch für den Vergleich der Frageblöcke untereinander.

3.1.3 Technische Ausstattung

Die Durchführung der Studie fand im Labor für experimentelle Kartographie der Universität Trier statt. Die verwendete Kamera zur Registrierung der Blickbewegung ist ein Tobii x50-Eyetracker der Firma Tobii aus Schweden. Wie bereits in Kapitel 2.3 beschrieben, verwendet dieses System die Cornea - Reflex - Methode zur Erfassung der Blickdaten. Hierzu besitzt das verwendete System zwei Nah - Infrarot - Dioden, die ein gebündeltes Feld Infrarotlichtstrahlen aussenden, sowie eine Videokamera, die den reflektierten Lichtstrahl und das Auge aufzeichnet [Ser07]. Die Kamera arbeitet hierbei mit einer Aufnahmefrequenz von 50 Hz, das heißt, dass alle 20 Millisekunden eine Messung stattfindet. Durch diese Aufnahmefrequenz ist es möglich, eine ausreichende Genauigkeit der Aufzeichnung zu gewährleisten, da ein Fixationsintervall nach Definition (siehe Kapitel 2.3) ab 50 Millisekunden beginnt. Es können also auch kürzeste Fixationen erfasst werden, was für den Kontext dieser Untersuchung im Hinblick auf den Zeitdruck während der Testphase von Vorteil ist.

Für die Verwendung der Blickbewegungskamera werden vom Hersteller folgende Mindest-Systemvoraussetzungen empfohlen [Tob05]:

- Intel Pentium 4, 2,4 GHz oder höher
- 265 MB Arbeitsspeicher
- Windows 2000 (SP 3 oder 4), Windows XP (SP 1 oder 2)

Aufgrund dieser geringen Hardwareanforderungen ist es möglich, auch komplexere Anwendungen auf dem Rechner ablaufen zu lassen, ohne mit der Eyetracking Software in einen Ressourcenkonflikt zu geraten. Bei zu hohen Anforderungen einer Anwendung ist zudem das Auslagern der Anwendung auf einen Zweitrechner möglich. Im Rahmen dieser Arbeit wurde allerdings der ressourcenschonende Slideshow Stimulus eingesetzt, wodurch kein weiteres System von Nöten war. Der Aufbau des

Systems sieht folgendermaßen aus: Am verwendeten Rechner sind zwei Monitore angeschlossen, ein Systemmonitor, von dem aus die Eyetracking Software gesteuert wird und ein Präsentationsmonitor, auf dem dem Probanden der entsprechende Stimulus, in diesem Fall eine Slideshow, gezeigt wird. Während des eigentlichen Tests ist es möglich, permanent am Systemmonitor zu prüfen, ob die Aufzeichnung der Blickbewegungsdaten korrekt funktioniert oder ob eine Neujustierung stattfinden muss. Den Aufbau des Systems, wie es im Labor für experimentelle Kartographie der Universität Trier zu finden ist, zeigt Abbildung 3.1.

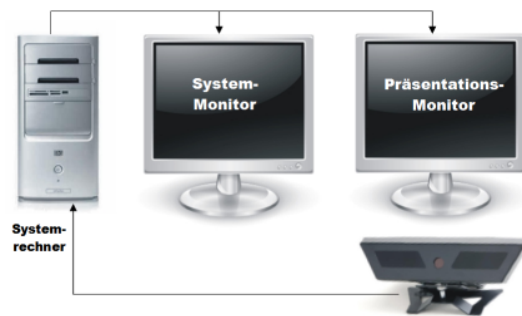


Abbildung 3.1: Systemanordnung der Studie

3.1.4 Ablauf der Studie

In diesem Unterkapitel wird der typische Verlauf einer Sitzung im Rahmen der Studie beschrieben. Die Gesamtdauer beträgt pro Proband ca. 45 Minuten, wobei dieser Wert inklusive der Vor- und Nachbereitung des Versuchs zu verstehen ist. Die einzelnen Durchläufe der Evaluation verteilen sich auf den Kernzeitraum von ca. einem Monat, wobei hier an einzelnen Tagen eine Laborzeit von über 10 Stunden zu verzeichnen war. Es erwies sich als schwierig, sowohl die Labor - Nutzungszeiten durch andere Projekte, als auch eine verbindliche Terminabsprache mit den Probanden miteinander in Einklang zu bringen.

Bei der Studie musste jeder Proband einzeln im Labor erscheinen, um eine Beeinflussung oder Störung durch Dritte, sowie einen vorweggenommenen Lerneffekt durch Beobachtung des Vorgängers, zu vermeiden. Bei Betreten des Labors wurde dem Probanden zunächst sein Sitzplatz zugewiesen. Im Anschluss daran erhielt er einen Fragekatalog, der alle Tasks enthält mit den exakten Formulierungen, wie sie auch im anschließenden Versuch auftreten. Dem Probanden wurde dann ausreichend Zeit eingeräumt, bis er sich den Katalog durchgelesen hat. Daraufhin

wurde durch den Versuchsleiter noch einmal der gesamte Katalog mit einer Reihe von ergänzenden Erklärungen durchgegangen. Insbesondere zu den Tasks 3 und 4 sind weitreichende ergänzende Anmerkungen von Nöten, da das Konzept eines Teilgraphen und einer Clique für den Großteil der Probanden nur schwer zugänglich sind, besonders im Hinblick darauf, dass ausschließlich nach der syntaktischen Information „Teilgraph“ oder „Clique“ zu suchen ist. Hier wurde den Probanden daher jeweils ein Beispiel für eine Clique und einen Teilgraph in allen drei zur Verwendung kommenden Layouts gezeigt, um auszuschließen, dass ein bestimmtes Layout von vorneherein einen Vorsprung gegenüber den anderen Layouts hat. Auch die Begriffe „maximaler Grad“ und „Weg“ beziehungsweise „Start-“ und „Zielknoten“ wurden jedem Probanden nochmals ausführlich erklärt. Ferner wurde bei jedem Task auf auftretende Besonderheiten hingewiesen, so z.B. im Fall von Task 5: Neben der Nennung des Knotens mit dem höchsten Grad, soll zusätzlich noch der Grad angegeben werden. Findet man zwei Kandidatenknoten mit gleichem Grad, so reicht es, einen davon anzugeben. Derart genaue Erklärungen waren nötig, da ein Großteil der Versuchsgruppe bisher keinerlei bewusste Erfahrungen mit Graphen und den zugehörigen Begrifflichkeiten hat und gleiche Startvoraussetzungen für jeden einzelnen Probanden gelten mussten. Nach dieser Einweisung wurden jedem Probanden noch einige Beispielgraphen gezeigt, die in den verschiedenen Layouts entworfen sind, um sicherzugehen, dass alle Layouts gleichermaßen als „korrekte Graphen“ anerkannt werden. Jeder Teilnehmer konnte dann noch auftretende Fragen stellen, so dass sämtliche Verständnisprobleme ausgeräumt waren.

Nach der Einweisung in die theoretischen Grundlagen des Versuchs, wurde dann jeder Teilnehmer über die konkreten Versuchsumstände informiert. Insbesondere wurde darauf hingewiesen, dass für die Lösung jeder Aufgabe ein Zeitlimit von 45 Sekunden vorgesehen ist. Zielvorgabe ist allerdings, die Aufgabe so korrekt und gleichzeitig schnell wie möglich zu beantworten. Sollten die 45 Sekunden überschritten werden, so wird zum nächsten Task übergegangen und eine möglicherweise gegebene Teillösung verworfen. Die Probanden mussten sich im Anschluss vor dem Präsentationsmonitor positionieren und wurden davon in Kenntnis gesetzt, dass sie diese Sitzposition während der Aufzeichnung nicht verändern dürfen, auch große Kopfbewegungen sollen vermieden werden. Alle Antworten sollen mündlich gegeben werden und können korrigiert werden, wenn noch keine endgültige Lösung abgegeben wurde. Eine endgültige Lösung ist die vollständige Beantwortung einer Aufgabe: Bei Task 1 die Aussage „ja“ oder „nein“, bei Task zwei ein Weg, der mit dem Zielknoten endet, bei Task 3 und 4 die Nennung der fünf bzw. vier beteiligten Knoten und bei Task 5 die Nennung eines Knotens und des zugehörigen Grads.

Bevor der eigentliche Versuch nun startete, wurde das Eyetracking System für jeden Probanden pro Frageblock kalibriert. Diese Kalibrierung wurde gespeichert und dem Profil des jeweiligen Probanden zugeordnet, auf diese Weise kann eine Kalibrierung wiederverwendet werden. Um allerdings kleine Veränderungen der Sitzposition und der Kopfneigung während der Pausen als Fehlerquelle auszuschließen, wurde vor jedem Frageblock eine Neukalibrierung vorgenommen. In der Kalibrierungsphase werden dem Probanden am Monitor Punkte auf einem weißen Hintergrund präsentiert, die er fixieren soll. So kann das System an die individuellen Blickbewegungen eines jeden Probanden angepasst werden. Nach erfolgreicher Kalibrierung folgte der erste Frageblock. Hier wird zunächst eine Frage eingeblendet, die sich konkret auf den folgenden Graph bezieht. Die Dauer dieser Einblendung bestimmt der Proband selbst, es ist hier kein Zeitlimit vorgegeben. Hat er die Frage verstanden, gibt er dies an und der Versuchsleiter blendet die nächste Folie ein. Wie bereits erwähnt, handelt es sich hierbei um ein Fixationskreuz, das für eine Dauer von exakt drei Sekunden eingeblendet wird und danach wird automatisch zur Folie mit dem Graph weitergeleitet. Dieser wird dann so lange gezeigt, bis der Proband eine vollständige Lösung im obigen Sinne angegeben hat, maximal jedoch 45 Sekunden. Unmittelbar nach der Lösungsangabe oder dem Ablauf des Zeitlimits wurde vom Versuchsleiter die nächste Frage eingeblendet.

Um eine übermäßige Belastung des Probanden zu vermeiden, wurde nach jedem Frageblock eine mehrminütige Pause eingelegt, deren Dauer der Proband selbst bestimmen konnte, allerdings war eine Mindestdauer von zwei Minuten vorgesehen. Während dieser Pause wurde der Proband angehalten, seinen Nacken zu entspannen, da durch die starre Haltung vor dem Präsentationsmonitor eine Verkrampfung, insbesondere der Nackenpartie, einhergeht. Ferner wurde der Proband in ein kurzes Gespräch verwickelt, um ihn von den gerade bewältigten Aufgaben abzulenken und so eine höhere Konzentrationsfähigkeit für die folgenden Frageblöcke zu erreichen. Nach Abschluss aller drei Frageblöcke wurde der Proband gebeten, einen Fragebogen zur anonymen Erfassung seiner Daten auszufüllen. Hierbei werden neben statistisch relevanten Informationen auch körperliche Merkmale und Verhaltensweisen abgefragt, die einen Einfluss auf die Ergebnisse haben können, so z.B. eine bekannte Einschränkung des Sehvermögens, Muttersprache oder auch die Dauer der Beschäftigung mit 3-D-Computerspielen. Auf diese Weise ist es möglich, Erklärungen für eventuell auftretende auffällige Einzelergebnisse zu finden.

Im Anschluss wurde jedem Probanden die Aufwandsentschädigung von 5€ausgezahlt. Es bleibt zu erwähnen, dass auch nach der recht anstrengenden Testprozedur ein großes Interesse an der Thematik seitens der meisten Probanden zu erkennen war. Dies wird durch eine Vielzahl von testbezogenen Fragen sowie Fragen über

die Hintergründe der Befragung deutlich. Dies zeigt, dass die Forschung auf diesem Gebiet durchaus fachbereichsübergreifend von großem Interesse ist.

3.2 Die Vorstudie

Zum Thema dieser Studie, nämlich dem Vergleich verschiedener Layoutverfahren mit Hilfe eines Eyetrackers, ist es sehr schwer, geeignete Vorlagen in der Literatur zu finden, da eine ähnliche Evaluation dieses Umfangs bisher nicht existiert. Es war also nicht möglich, auf Erfahrungswerte bereits abgeschlossener Untersuchungen zurückzugreifen. Aus diesem Grund wurde der eigentlichen Evaluation eine Studie vorangeschoben, die die Aufgabe hatte, als eine Machbarkeitsstudie zur Hauptevaluation zu dienen. Diese Vorstudie sollte sich mit dem gleichen Thema befassen wie die spätere Evaluation, so dass sich einerseits Fehler in der Methodik herauskristallisieren und auf der anderen Seite Aussagen über den Umfang und Schwierigkeitsgrad der Aufgaben und Graphen getroffen werden konnten. Ferner sollten erste Erfahrungen mit dem Eyetracking System gemacht werden und dabei insbesondere der Output der Software daraufhin überprüft werden, inwiefern dieser für eine fundierte Analyse zu gebrauchen ist und in welcher Form. Ein vorher unbeabsichtigter Nebeneffekt der Vorstudie war es, dass sich bei der quantitativen Auswertung der Ergebnisse bereits sehr interessante Thesen und eindeutige Tendenzen gezeigt haben, die nun als Hypothesen in der Hauptstudie auf statistisch feste Beine gestellt werden sollen.

Zur Vorstudie wurden fast alle Vorbereitungen getroffen, wie sie auch bei der Hauptstudie vorkommen. Die Tasks mussten entworfen werden, so dass ein möglichst breites Feld der Probleme der Graphentheorie abgedeckt und zugleich die Komplexität für eine Evaluation nicht zu hoch angesetzt war. Hier wurden bereits die oben aufgeführten Tasks festgelegt, wie sie auch in der Hauptstudie später vorkommen sollen. Auch die Reihenfolge der Tasks und die Formulierung der Fragen entsprach bereits weitestgehend der Version in der Hauptevaluation. Die in der Hauptstudie verwendeten Graphen entsprechen zu einem Großteil ebenfalls bereits denen, die für die Vorstudie entworfen wurden. Hier gab es noch einige Verbesserungen vorzunehmen, auf die später noch eingegangen wird.

Der Ablauf der Vorstudie entsprach in kleinerem Maßstab dem der späteren Evaluation. Ein Hauptunterschied zwischen beiden Studien stellt die Größe der Versuchsgruppe dar. An der Vorstudie nahmen nur fünf Probanden teil, anhand derer

die hier benötigten Erfahrungen für die Evaluation gesammelt wurden. Die Zusammensetzung der Probandengruppe entsprach allerdings bereits in Grundzügen der Zusammensetzung der zu erwartenden Teilnehmergruppe an der Hauptevaluation. Auch hier beschränkte sich die Auswahl an Kandidaten nicht nur auf Informatikstudenten, sondern zusätzlich nahmen auch Geographen und Wirtschaftswissenschaftler teil. Dadurch konnte bereits hier festgestellt werden, dass der Schwierigkeitsgrad der Fragen auch für themafremde Studenten nicht zu hoch ist. Hier konnten wichtige Erkenntnisse für die spätere Evaluation gemacht werden, was die Erklärung der theoretischen Hintergründe anbelangt. Insbesondere die Beschreibung der Eigenschaften „Teilgraph“ und „Clique“ entwickelte sich im Verlauf der Vorstudie zu der Form, wie sie letztendlich auch in der Hauptstudie den Probanden nahe gebracht wird.

Auch die Fähigkeiten des Eyetracking Systems, mit Brillen- oder Kontaktlinsenträgern umzugehen, wurde in der Vorstudie durch die entsprechende Auswahl an Probanden getestet. Hier zerstreuten sich die Befürchtungen, das System könnte bei derartigen Sehhilfen Probleme bei der Aufzeichnung der Blickbewegung bekommen. Im Gegenteil zeigt sich eine hohe Toleranz der Kamera gegenüber Sehhilfen aller Art.

Vergleicht man die Graphen, die bei Vor- und Hauptstudie als Grundlage dienen, so kann man feststellen, dass der grundsätzliche Aufbau und die syntaktischen Informationen bei sich entsprechenden Graphen gleich sind. Man könnte sagen, die Graphen sind isomorph. Es gibt nur wenige Unterschiede in der Darstellung der Graphen in beiden Studien. Ein gravierender Unterschied ist die Benennung der Knoten. Hier wurden bei allen Graphen der Vorstudie zwar ebenfalls Großbuchstaben des Alphabets als Labels gewählt, allerdings fingen die Knotennamen bei „A“ an und setzten sich entsprechend ihrer Reihenfolge im Alphabet fort. In der Vorstudie wurde schnell deutlich, dass eine solche Benennung suboptimal ist. Da die Probanden die Lösung der jeweiligen Aufgabe immer mündlich formulieren mussten, kam es häufig zu Verständnisschwierigkeiten aufgrund von phonetisch ähnlichen Buchstabenkombinationen. Die Konsequenz für die Hauptstudie ist, dass dort eine Auswahl von phonetisch möglichst klar unterscheidbaren Buchstaben getroffen wird, anstatt einfach von „A“ beginnend aufsteigend zu benennen. Hierbei konnte eine vollständige phonetische Unabhängigkeit nicht erreicht werden, da bei einem Graph, der aus 20 Knoten besteht, natürlich auch 20 Buchstaben zur Benennung gebraucht werden, was bei insgesamt 26 möglichen Buchstaben zu Kompromissen bei der Auswahl führt. Dennoch konnte in der Hauptstudie eine deutliche Verbesserung der Verständlichkeit der mündlich abgegebenen Antworten erreicht werden. Hier hat sich die Vorstudie also ebenfalls bewährt.

Eine weitere Verbesserung der Graphdarstellung schien nach Ende der Vorstudie ebenfalls für Graphen im hierarchischen Layout von Nöten. Der Grund lag darin, dass das hierarchische Layout in fast allen Disziplinen deutlich hinter den Ergebnissen der anderen beiden Layoutverfahren zurücklag. Dies könnte natürlich tatsächlich darauf hinweisen, dass das hierarchische Layout überproportional große Nachteile gegenüber den anderen Verfahren hat, bei einer genaueren Betrachtung äquivalenter Graphen in verschiedenen Layoutvariationen fiel jedoch besonders der Umstand auf, dass der vertikale Abstand der einzelnen Schichten sehr gering gewählt war. Dieser entsprach der Standardeinstellung von DGD, das für das Layouten verantwortlich zeichnet. Als Konsequenz aus der sehr engen Schichtung der Hierarchieebenen verliefen die Kanten (bzw. Bögen) zum Teil fast waagrecht aus einem Knoten heraus, machten dann einen scharfen Bogen und liefen ähnlich waagrecht zum Zielknoten. Die gesamte Abbildung wirkte in vertikaler Richtung gestaucht. Durch diese Stauchung und den damit verbundenen teils recht waagerechten Verlauf vieler Kanten kam es bei Kantenkreuzungen in sehr vielen Fällen zu sehr spitzen Winkeln, was die Lesbarkeit der Graphen offensichtlich stark einschränkte. Da es sich hier weniger um ein Layout-spezifisches Problem handelt, sondern viel mehr eine Parameteranpassung in DGD von Nöten war, um hier eine Gleichstellung der drei Layoutverfahren zu gewährleisten, wurden die hierarchischen Graphen in der Hauptstudie mit einem vertikalen Abstand versehen, der etwa dem vertikalen Abstand entspricht, der beim orthogonalen Verfahren zur Verwendung kommt. Ob diese Verbesserung tatsächlich große Auswirkungen auf die Studie haben wird, wird später ausführlich betrachtet.

Eine weitere grundsätzliche Frage, die im Rahmen der Vorstudie zu klären war, ist die Anordnung der Fragen in den einzelnen Frageblöcken. Von vornherein stand fest, dass die Verwendung eines Stimulus pro Frage keinen Sinn machen würde, da dadurch die Slideshow zu sehr in einzelne Bestandteile zerlegt würde. Man kann leicht erkennen, dass es dadurch für den Probanden fast unmöglich ist, eine dauerhafte Konzentration aufrecht zu erhalten, da nach jeder beantworteten Frage der Versuch kurz unterbrochen werden würde. Außerdem würde eine derartige Vielfalt von Stimuli zu einem starken Anstieg des Verwaltungs- und Auswertungsaufwandes führen. Auf der anderen Seite dürfen auch nicht zu viele Fragen zu einem Block zusammengefasst werden, da dadurch keine Pausen für den Probanden mehr möglich sind. Durch die bereits angesprochene starke Anspannung für die Versuchspersonen, insbesondere für deren Nackenmuskulatur, aber auch für die Konzentrationsfähigkeit, sind regelmäßige Pausen zwingend notwendig. Ziel der Vorstudie war also auch, diesen Trade-off zu optimieren, was durch die Verwendung von verschiedenen großen Frageblöcken realisiert wurde. Es kamen Frageblöcke zum Einsatz, die 5, 10, 15 und 20 Fragen in sich vereinten. Hierbei zeigt sich sehr schnell, dass die kurzen Blö-

cke mit 5 oder 10 Fragen von den Versuchspersonen als zu kurz empfunden wurden („Schon vorbei...?“ und ähnliche Äußerungen wurden hier unaufgefordert geäußert). Es entstand der Eindruck, dass gerade beim Höhepunkt der Konzentrationsphase der Block endet. Entgegengesetzte Reaktionen gab es bei einer Frageblockgröße von 20 Fragen. Hier wirkten die Probanden bei den letzten Fragen unkonzentriert und gleichgültig, außerdem nahmen die Kopfbewegungen stark zu, wohl um die Anspannungen im Nackenbereich zu lockern. Nachdem der Block dann abgeschlossen war, merkte man den Probanden die Erleichterung an. Als optimale Größe eines Frageblocks erwies sich daher eine Komplexität von 15 Fragen. Die Bearbeitungsdauer für einen solchen Frageblock lag mit ca. 5-8 Minuten bei einem Wert, der es ermöglichte, über die gesamte Dauer hinweg konzentriert zu arbeiten und der gleichzeitig körperlich nicht zu anstrengend ist.

Doch nicht nur die optimale Blockgröße galt es zu finden, ein weiterer Designaspekt ist die Reihenfolge der verschiedenen komplexen Graphen in den jeweiligen Layouts. Die Reihenfolge der Zuordnung der Fragen zu den Graphen war in der Vorstudie nicht zufällig, sondern folgte einem System. Es wurden immer alle fünf Tasks hintereinander zu der gleichen Abbildung abgearbeitet. Das heißt der erste Taskdurchlauf bezog sich immer auf den kräftebasierten 10-Knoten Graph, der nächste Durchlauf betrachtete immer den orthogonalen 10-Knoten Graph und der dritte Durchlauf den hierarchischen 10-Knoten Graph. Anschließend folgten die 15-Knoten Graphen und die 20-Knoten Graphen in der gleichen Layoutreihenfolge. Dieses Verfahren erwies sich als nicht praktikabel und konnte auch keine Lerneffekte verhindern. Es treten unweigerlich Übungseffekte ein, wenn ein Task zunächst für einfache Graphen mehrfach bearbeitet und die Komplexität der Graphen dann langsam gesteigert wird. Eine unabhängige quantitative Aussage kann dann insbesondere für die komplexesten Graphen nicht abgegeben werden, da deren Ergebnisse von Lerneffekten beeinflusst sind. Als Konsequenz aus dieser Erkenntnis wurden bei der Hauptstudie sowohl die Reihenfolge der Layouts, als auch die Reihenfolge der Komplexität der Graphen zufällig permutiert, so dass kein Layout und kein Graph einen Vorteil aus einer festgelegten systematischen Reihenfolge ziehen kann.

Eine weitere Maßnahme zur Minimierung von Übungs- bzw. Lerneffekten, die nach der Auswertung der Vorstudie getroffen wurde, ist die Permutation der Knotennamen bei gleich komplexen Graphen in unterschiedlichen Layouts. Die gesamte Studie basiert auf drei unterschiedlich komplexen Graphen (10 Knoten, 15 Knoten, 20 Knoten), die jeweils in drei verschiedenen Layoutverfahren (kräftebasiert, orthogonal, hierarchisch) gestaltet wurden. Bei der Vorstudie wurden dabei sich entsprechende Knoten in den drei verschiedenen Layouts mit dem gleichen Großbuchstaben versehen. Dies erleichtert erheblich den Auswertungsaufwand, da auf diese Weise

ein wesentlich kompakterer und vor allem layoutunabhängiger Lösungskatalog zur Kontrolle erstellt werden kann. Der Nachteil dieser Benennung zeigt sich jedoch im Laufe der Vorstudie recht schnell, da die Probanden gewisse immer wiederkehrende Lösungen (da ja zu jedem Task drei gleiche Lösungen abgegeben werden konnten) von vorneherein präferierten. Dies äußerte sich in Aussagen wie „Die Lösung ist *wieder XYZ*.“ oder „*Wie gerade eben*: Knoten X mit Grad Y.“. Es ist zwar anzunehmen, dass die Probanden nicht nach bestimmten Knotenlabels gesucht haben und anschließend diese auf die gefragten Eigenschaften überprüft haben, insbesondere da kein Proband die verschieden layouteten Graphen gleicher Komplexität miteinander in eine isomorphe Verbindung brachte, um einen möglichen Lerneffekt aber trotzdem auszuschließen, wurden die Knotennamen sich entsprechender Knoten gleich komplexer Graphen zufällig permutiert. Der Namensraum umfasst zwar die gleichen Buchstaben, allerdings sind sich entsprechende Knoten stets anders benannt.

Insgesamt betrachtet lieferte die Vorstudie ganz entscheidende Erfahrungswerte, die einen sehr großen Anteil am Erfolg der Hauptstudie haben. Durch die wohlüberlegte Planung der Vorstudie konnten bereits eine Vielzahl von Rahmenbedingungen so vorgefunden werden, wie sie auch in der eigentlichen Evaluation auftreten. Das Zeitlimit von 45 Sekunden erwies sich als optimal, denn dadurch entstand ein gewisser Zeitdruck für die Probanden, was zur Intensivierung der Konzentration führte. Außerdem kann ein Kandidat bei unbeschränkter Zeit so lange eine Lösung suchen, bis er eine optimale Antwort gefunden hat. Dies ist nicht der Sinn dieser Evaluation, denn es geht darum zu überprüfen, welches Layout besonders gut für die Unterstützung verschiedener Tasks geeignet ist. Bei beliebiger Zeit könnte daher ein impliziter Vorteil eines bestimmten Layouts kaum mehr nachgewiesen werden. Die time-out - Quote bei der gesamten Vorstudie lag bei nur 3%, das zeigt, dass auch der Wert von 45 Sekunden sehr gut gewählt ist. Betrachtet man nämlich die Bearbeitungszeiten der einzelnen Tasks, so kommt man zu dem Schluss, dass bei einer Kürzung des Zeitlimits auf 30 Sekunden die time-out - Quote insbesondere bei anspruchsvollen Tasks und komplexen Graphen so stark ansteigt, dass verwertbare Ergebnisse nicht mehr gegeben sind. Auch die Größe der Graphen wurde von vornherein treffend gewählt, denn während die 10-Knoten Graphen von allen Probanden mit einer sehr hohen Erfolgsquote und in der Regel auch weit unter dem Zeitlimit gelöst wurden, mehrten sich die Fehler rapide, wenn die 20-Knoten Graphen bearbeitet wurden. Auch die time-out - Quote stieg hier stark an. Dies zeigt, dass bei einem Graph mit weniger als 10 Knoten keine sonderliche Vereinfachung mehr zu erwarten ist, da bereits der 10-Knoten Graph von allen Probanden als sehr leicht zugänglich empfunden wurde. Eine Erhöhung der Knotenkomplexität auf über 20 Knoten (mit entsprechender Erhöhung der Kantenzahl, um den Durchschnittsgrad von ca. 3 zu erhalten) würde Fehler- und time-out - Quote wohl in Bereiche treiben, in denen

sinnvolle Auswertungen wohl nicht mehr möglich sind. Denn bereits zwischen den 15-Knoten und den 20-Knoten Graphen ist ein starker Anstieg der falschen Antworten und der Bearbeitungszeiten zu erkennen. Auch die Tasks können weitestgehend unverändert aus der Vorstudie übernommen werden. Die Mischung zwischen *subjektiv* einfachen Tasks (Task 1), mittelschweren Tasks (Tasks 2 und 5) und schweren Tasks (Tasks 3 und 4) erwies sich als praktikabel. Durch diese Abstufung konnte die Vorstudie für die Probanden abwechslungsreich gestaltet werden, was wiederum zur Steigerung der Konzentration beiträgt. Außerdem konnten durch die Tasks einige wichtige graph- bzw. komplexitätstheoretische Probleme abgedeckt werden, wodurch ein entsprechender Nutzen für diese Fachgebiete gewonnen werden kann.

Neben dem Hauptzweck der Vorstudie, nämlich dem Gewinnen von Erfahrungswerten für die Hauptevaluation, ergaben sich als ein Nebenprodukt einige sehr interessante Thesen, die einer Belegung oder Widerlegung durch die Hauptstudie bedürfen. An dieser Stelle sollen nur die wichtigsten Tendenzen angeführt werden, die es in der Hauptevaluation zu überprüfen gilt. Eine komplette Analyse der Ergebnisse würde den Rahmen der Arbeit sprengen und zudem keine statistisch belegbaren Ergebnisse liefern, da die Probandengruppe nicht die ausreichende Größe hatte. Eine intensive quantitative Analyse der Ergebnisse der Hauptstudie wird in Kapitel 3.3 zu finden sein. Eine erste Tendenz, welche Layoutvarianten für einen menschlichen Betrachter am zugänglichsten ist, kann man im Rahmen der Vorstudie erhalten, wenn man die Erfolgsquoten der einzelnen Tasks bei den verschiedenen layouteten Graphen betrachtet. In Abbildung 3.2 bis 3.5 sind die Gesamterfolgsquoten der Tasks 2 bis 5 abgebildet. Dabei wird lediglich beachtet, ob die gegebene Lösung korrekt oder falsch ist oder ob eine Überschreitung des Zeitfensters von 45 Sekunden vorliegt.

Betrachtet man die Ergebnisse der unterschiedlichen Layouts bei Task 2, so fällt auf, dass das kräftebasierte Layout ein gutes Ergebnis erreichen konnte mit 87% korrekt beantworteten Fragen. Dies entspricht dem erwarteten guten Ergebnis, denn insbesondere das kräftebasierte Layout gilt als leicht zugänglich, da es auf physikalischen Begebenheiten beruht, die dem Menschen in der Natur tagtäglich begegnen. Allerdings stellt sich natürlich die Frage, wo bei 13% falschen Antworten die Fehlerquellen liegen. Sollte sich dieses Ergebnis in der Hauptstudie also bestätigen, so wird es interessant sein, mittels der Eyetrackerdaten zu analysieren, wo hier genau die Fehlerquellen liegen und welche Konsequenzen sich darauf für eine Modifikation des Verfahrens ergeben. Die eigentliche Überraschung bei Task 2 stellt allerdings das orthogonale Layout dar. Hier waren alle Probanden in der Lage, einen korrekten Weg vom Start- zum Zielknoten anzugeben und damit konnte das Ergebnis des kräftebasierten Layouts noch übertroffen werden. Sollte sich diese Tendenz statistisch belegen lassen, so müsste also bei ähnlich gelagerten Problemen die Nutzung eines

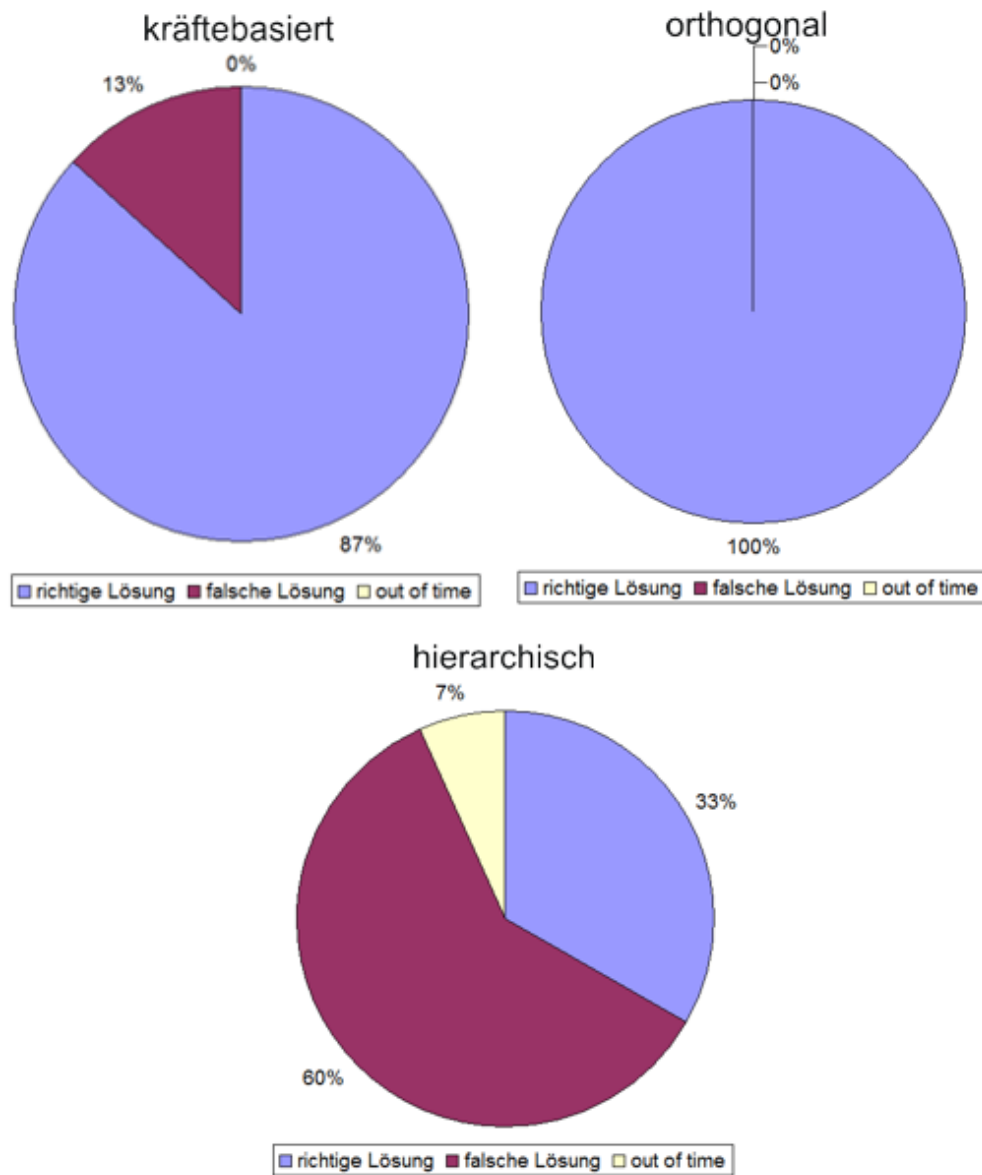


Abbildung 3.2: Gesamterfolgsquoten Task 2

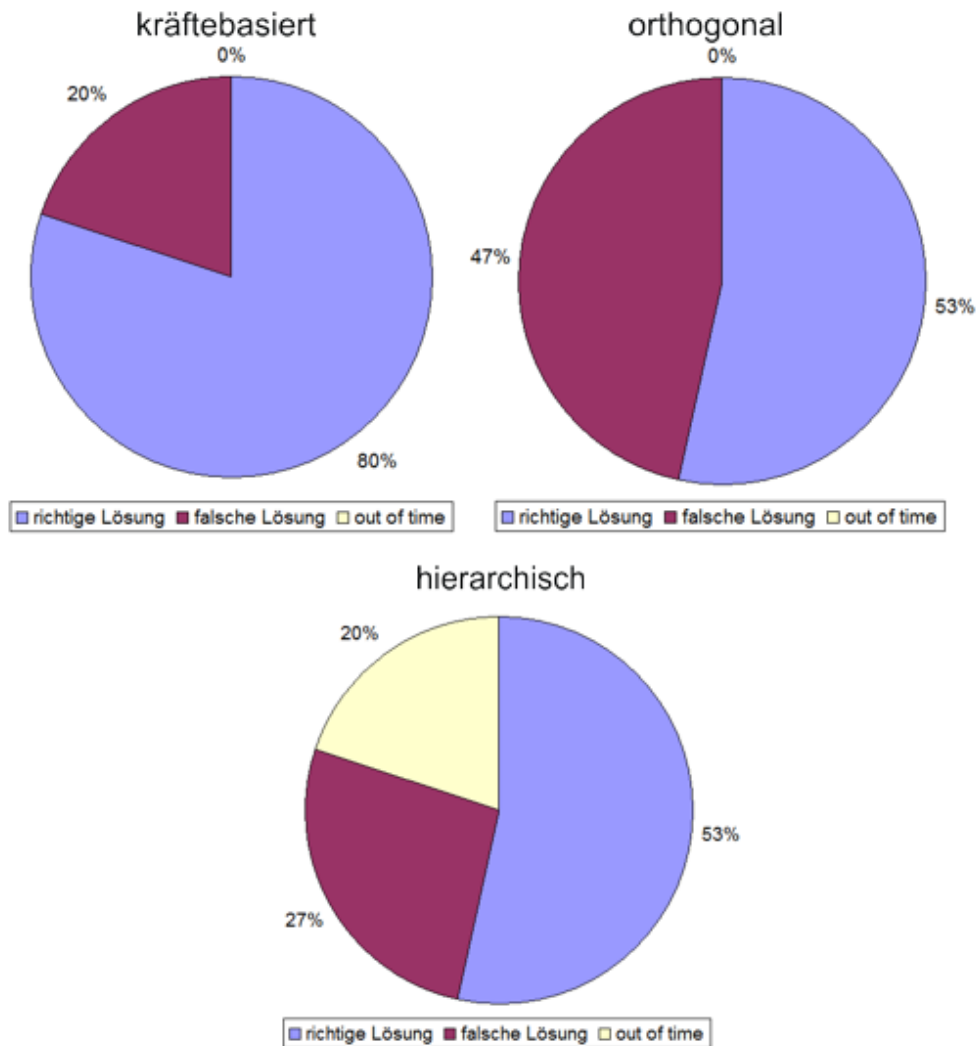


Abbildung 3.3: Gesamterfolgsquoten Task 3

orthogonalen Layouts einen entscheidenden Vorteil gegenüber dem kräftebasierten Layout haben. Betrachtet man nun noch die Ergebnisse des hierarchischen Layouts bei Task 2, so kann man feststellen, dass dieses gegenüber den beiden anderen Verfahren deutlich schlechter abschneidet mit nur 33% korrekten Antworten. Hier gilt aber zu berücksichtigen, dass die oben aufgeführten Darstellungsmängel einen großen Anteil an diesem schlechten Ergebnis haben können. Es wird sich also in der Hauptevaluation zeigen, ob die Korrektur des vertikalen Abstands der Schichten untereinander einen großen Einfluss auf das Ergebnis hat oder ob das hierarchische

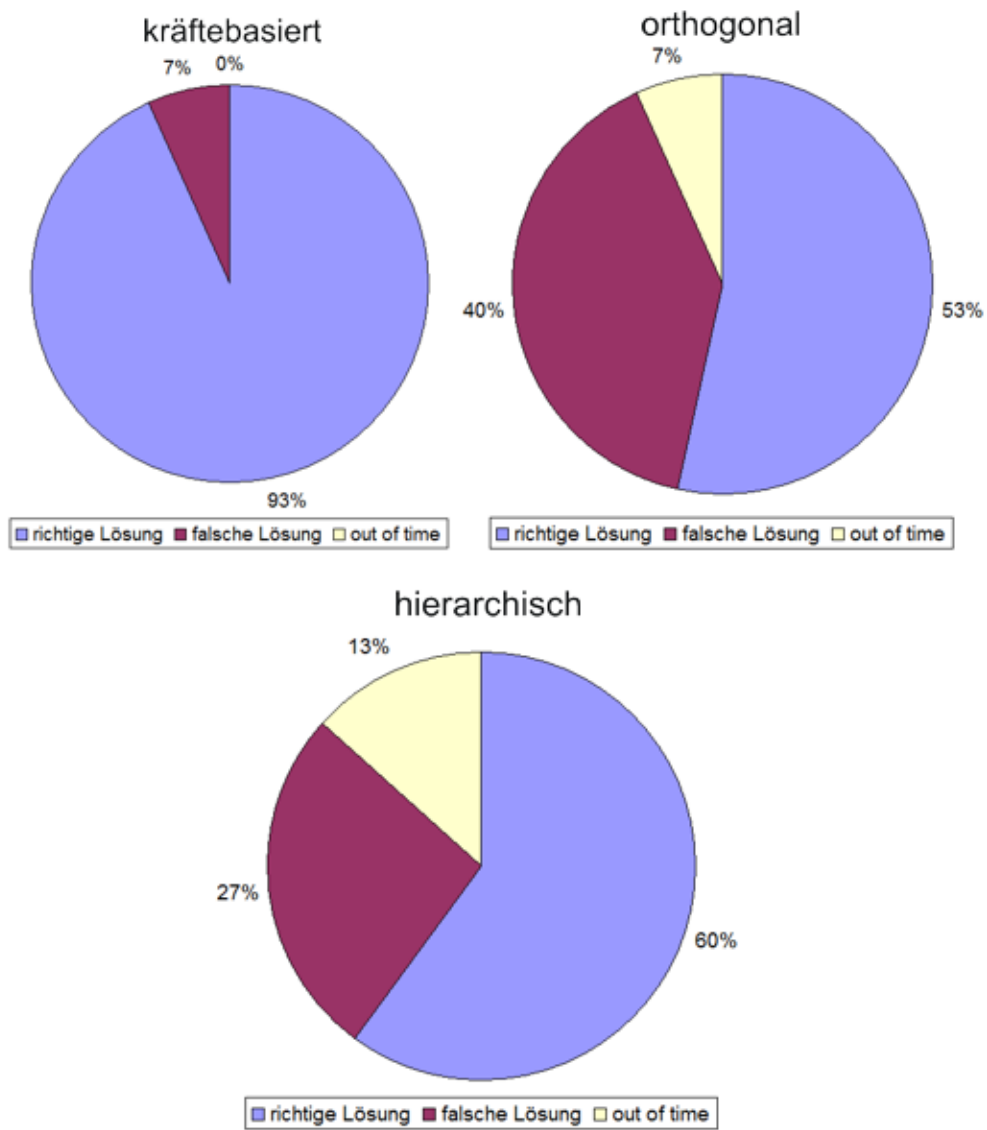


Abbildung 3.4: Gesamterfolgsquoten Task 4

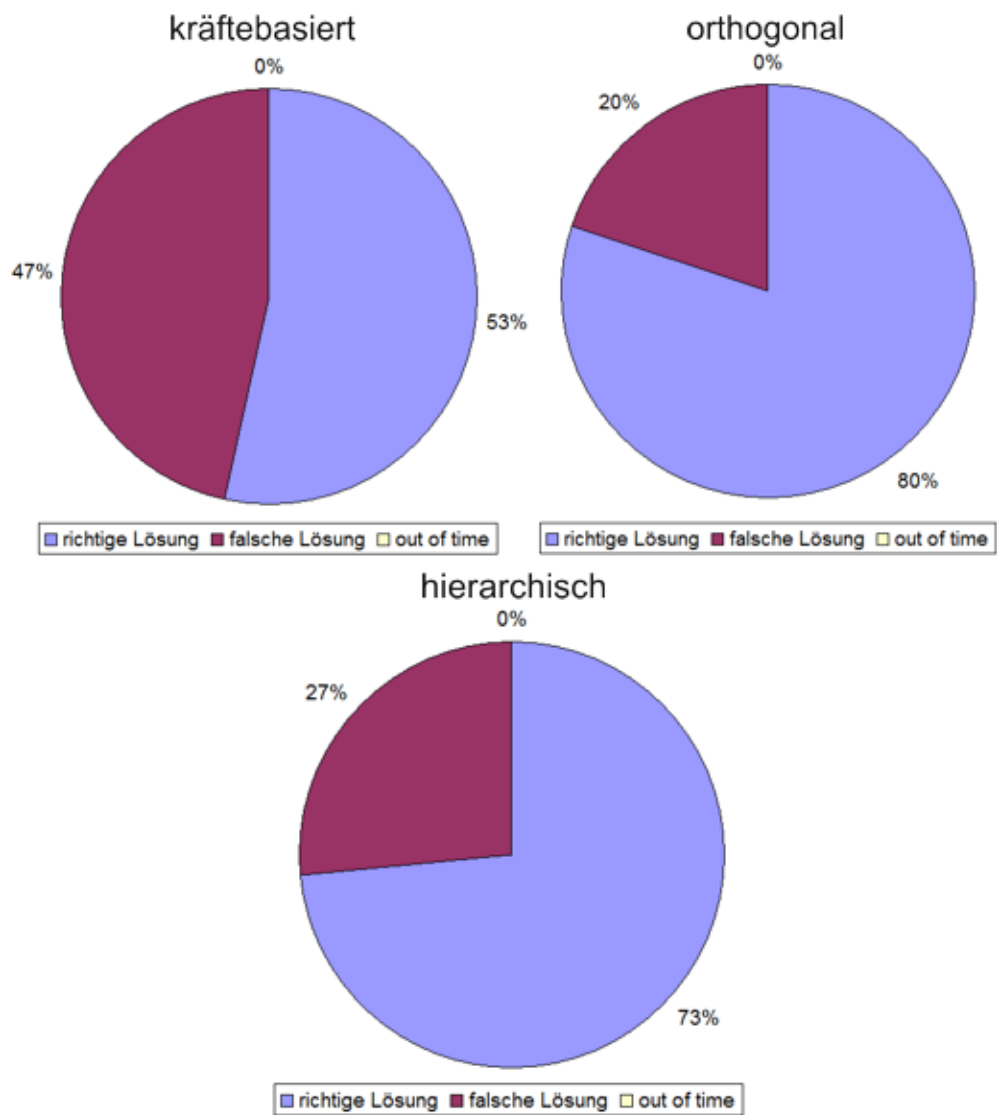


Abbildung 3.5: Gesamterfolgsquoten Task 5

Layout tatsächlich für Probleme, die ähnlich wie Task 2 gelagert sind, ungeeignet ist.

Bei der Betrachtung der Ergebnisse der Vorstudie fassen wir nun die Tasks 3 und 4 (Abbildungen 3.3 und 3.4) zusammen, da deren Ergebnisse ähnlich ausfallen und, wie bei der Beschreibung der Tasks bereits angedeutet, eine gewisse Verwandtschaft beider Tasks vorhanden ist. In der späteren Analyse der Hauptstudie wird jeder Task separat betrachtet werden, hier sollen jedoch nur Tendenzen erkannt werden, die es später zu überprüfen gilt. Bei den Tasks 3 und 4 scheint es eine eindeutige Präferenz der Probanden für das kräftebasierte Layout zu geben. Dieses konnte mit einem Anteil von 80% bzw. 93% korrekter Antworten die beiden anderen Verfahren deutlich auf die Ränge verweisen. Ein gutes Ergebnis dieses Verfahrens war insbesondere bei Task 4 zu erwarten, da eine Clique in der Regel eine sehr typische kräftebasierte Darstellung hat und die Probanden so (unbewusst) nach einer „Form“ suchen konnten, ähnlich ist der Erfolg bei Task 3 zu erklären. Dies wird in der Hauptstudie zu überprüfen sein. Beim Vergleich des orthogonalen mit dem hierarchischen Ergebnis überrascht, dass das hierarchische Layout bei Task 4 noch etwas besser zugänglich scheint, während Task 3 beide Verfahren exakt gleichauf sieht. Hier steht die Theorie im Raum, dass das orthogonale Layout die „Form“ einer bestimmten Knotenkonstellation in der Abbildung fast völlig zerstört durch überlange Kanten und rechtwinklige Knicke der Kanten. Das hierarchische Layout erhält durch die Vorgabe, lange Kanten nach Möglichkeit zu vermeiden und einen fließenderen Kantenverlauf eine gewisse äußere „Form“ aufrecht. Die Hauptevaluation wird zeigen, ob sich diese Thesen bestätigen lassen.

Task 5 liefert beim Layoutvergleich in der Vorstudie ebenfalls überraschende Ergebnisse. So kann sich hier das orthogonale Layoutverfahren an die Spitze setzen, was den Anteil an korrekt abgegebenen Antworten angeht. Scheinbar eignet sich eine rechtwinklige Anordnung von Kanten besonders gut, um einen Knoten mit hohem Grad schnell zu finden. Dies würde bedeuten, dass das orthogonale Verfahren bei Aufgaben Verwendung finden sollte, die als Kriterium den Knotengrad berücksichtigen, z.B. bei der Analyse der Zentralität von Knoten als Indikator für Macht in einem sozialen Netzwerk. Aber auch das hierarchische Layout kann in der Vorstudie bei diesem Task überzeugen, denn es erreicht mit 73% ebenfalls ein gutes Ergebnis. Dies könnte damit zusammenhängen, dass der Algorithmus einen Knoten mit vielen ausgehenden Kanten (und damit vielen Nachbarknoten) meist recht zentral in der Abbildung anordnet, wo er dem Betrachter deutlicher auffällt. Es wird interessant sein zu prüfen, ob die Modifikationen an den hierarchischen Graphen dieses Ergebnis noch verbessern, oder ob es im Hinblick auf Task 5 sogar eine Verschlechterung gibt. Ein sehr schlechtes Ergebnis konnte bei diesem Task das kräftebasierte Layout

erreichen. Mit nur 53% korrekter Antworten scheint dieses Verfahren sich für eine Aufgabe, die den Knotengrad zum Thema hat, nicht zu empfehlen. Allerdings muss auch hier die Hauptstudie zeigen, ob dies als ein zufälliges Ergebnis angesehen werden muss oder ob sich diese Schwachstelle bestätigt. Eine weitere Charakteristik für den Vergleich der Layoutverfahren liefert die Übersicht über die durchschnittlichen Bearbeitungszeiten pro Task (vgl. die Abbildungen 3.6 bis 3.9).

Abbildung 3.6 liefert eine Übersicht darüber, wie lange die Probanden durchschnittlich für die Bearbeitung der neun verschiedenen Abbildungen gebraucht haben. Ist bei einem Graph kein Wert eingetragen, so hat dies den Grund, dass alle Probanden die betreffende Frage falsch beantwortet haben. Dies ist bei Task 2 der Fall, denn keiner der Probanden konnte die Aufgabe für den hierarchischen 15-Knoten und 20-Knoten Graph lösen. Dieser Nachteil des hierarchischen Layouts wurde bereits bei der Betrachtung der Erfolgsquoten (Abbildung 3.2) deutlich und wird hier bestätigt. Auffallend bei dem Vergleich der Layouts ist, dass bei den 10-Knoten Graphen das hierarchische Layout im Gegensatz dazu sogar die kürzeste Bearbeitungszeit hat. Ob hier ein Lerneffekt die Ursache ist (der durch verschiedene oben genannte Maßnahmen bei der Hauptstudie minimiert wird) oder ob es bei Graphen mit einer geringen Komplexität tatsächlich einen Zeitvorteil für das hierarchische Layout gibt, wird die Hauptstudie zeigen. Eine Tendenz, die es ebenfalls zu untersuchen gilt, ist,

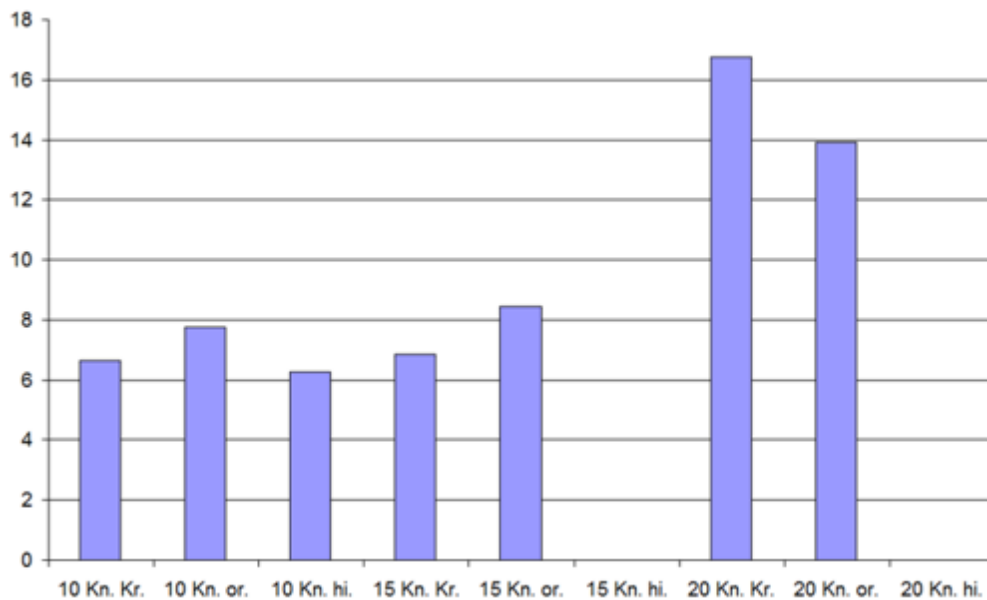


Abbildung 3.6: Durchschnittliche Bearbeitungszeit in Sek. (Task 2)

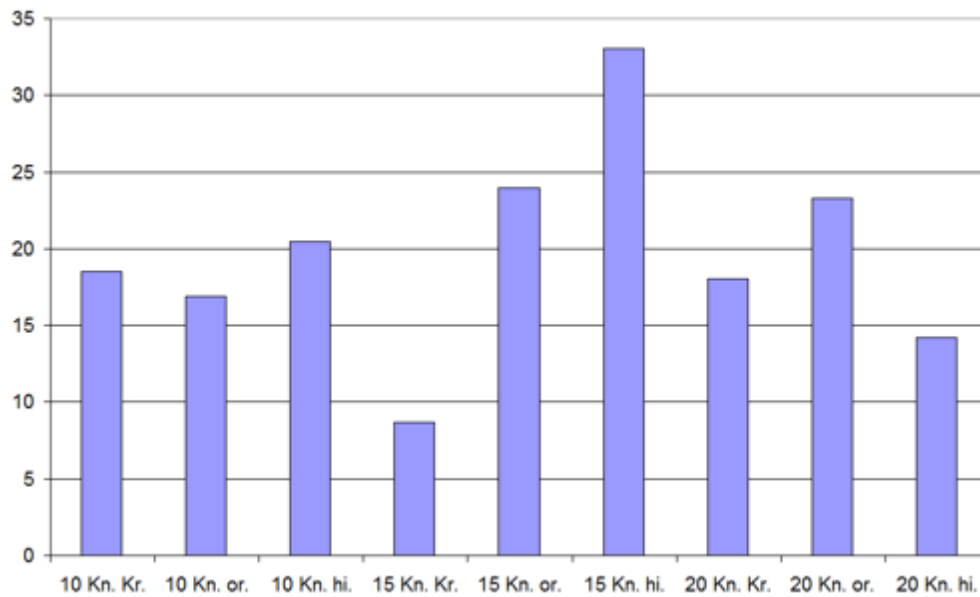


Abbildung 3.7: Durchschnittliche Bearbeitungszeit in Sek. (Task 3)

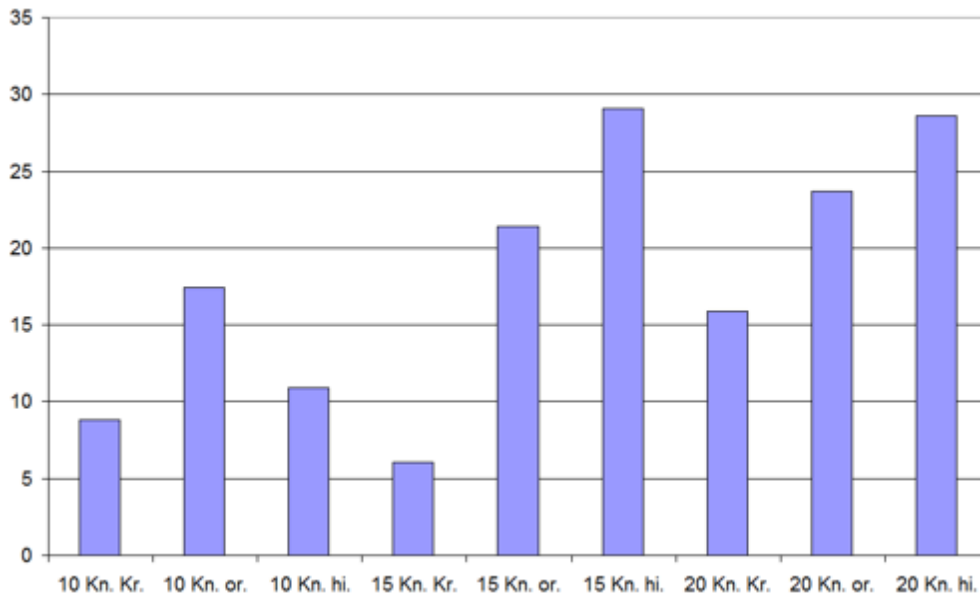


Abbildung 3.8: Durchschnittliche Bearbeitungszeit in Sek. (Task 4)

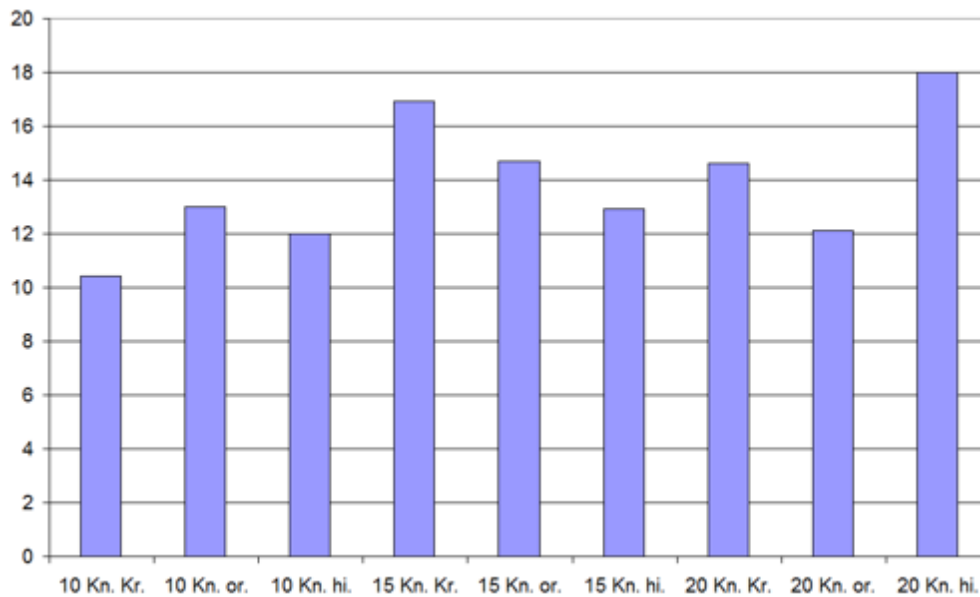


Abbildung 3.9: Durchschnittliche Bearbeitungszeit in Sek. (Task 5)

dass das kräftebasierte und das orthogonale Layoutverfahren bei Graphen gleicher Komplexität kaum Unterschiede in der Bearbeitungszeit aufweisen. Bei komplexeren Graphen kann das orthogonale Verfahren sogar eine kürzere Zeit aufweisen, was wiederum die These bestätigen würde, dass das Problem, einen Weg im Graphen zu finden, durch eine orthogonale Graphdarstellung besonders zugänglich wird. Auch hier soll die Hauptevaluation eine eindeutigere Tendenz aufweisen.

Betrachtet man die Abbildungen 3.7 und 3.8, so kann man feststellen, dass es eine deutliche Tendenz zu geben scheint, dass besonders bei komplexeren Graphen das kräftebasierte Layout von den Probanden am schnellsten bearbeitet werden kann, das orthogonale Layout nimmt einen Mittelplatz ein und die hierarchischen Darstellungen benötigen die meiste Bearbeitungszeit. Hierbei ist zu beachten, dass die niedrige Bearbeitungszeit von Task 3 bei der 20-Knoten hierarchischen Darstellung so zu erklären ist, dass hier nur eine einzige Lösung in die Durchschnittszeit eingeht, da alle anderen Probanden diese Frage falsch beantwortet haben. Daher muss dieser Wert bei der Betrachtung ausgeklammert werden. Eine weitere Auffälligkeit ist, dass die Zeitspanne über alle Graphen hinweg beim orthogonalen Layout wesentlich konstanter bleibt, als es bei den beiden anderen Verfahren der Fall ist. Hier muss die Hauptevaluation zeigen, ob sich dieser Trend bestätigt und wenn dies so ist, geben die Eyetrackerdaten eine Information für die Gründe dieses Verhaltens? Wie bei allen Analysen wird auch ein besonderes Augenmerk auf die Veränderung der

Werte für das hierarchische Layout gelegt, denn besonders für die Tasks 3 und 4 sind durch die Entzerrung der Graphen in der Hauptstudie bessere Ergebnisse zu erwarten.

Eine These, die durch Abbildung 3.9 in der Hauptstudie in den Fokus der Betrachtung rücken wird, ist, dass Task 5 kaum von der Komplexität der Graphen beeinflusst wird und dass auch die Layouts auf die Bearbeitungszeit (im Gegensatz zur Bearbeitungsqualität) nur einen geringen Einfluss haben. Vergleicht man jeweils Graphen gleicher Komplexität, so weisen der *orthogonale* 10-Knoten Graph, der *kräftebasierte* 15-Knoten Graph und der *hierarchische* 20-Knoten Graph jeweils die längsten Bearbeitungszeiten auf. Es lässt sich also keine klare Tendenz zu einem bestimmten Layout feststellen. Auch die Zeitspanne zwischen der schnellsten und der langsamsten Bearbeitungszeit über alle neun Graphen weist keine signifikante Entwicklung mit steigender Komplexität auf. Insbesondere zwischen den 15-Knoten und den 20-Knoten Graphen lässt sich keine Verlangsamung feststellen und auch die Entwicklung zwischen 10-Knoten und 15-Knoten ist nur sehr gering ausgefallen. Ob die Layouts tatsächlich nur einen Einfluss auf die Qualität der Antwort haben, die Bearbeitungszeiten allerdings kaum beeinflussen, soll daher die Analyse des entsprechenden Tasks in der Hauptstudie zeigen.

3.3 Quantitative Analyse

In der nun folgenden quantitativen Analyse der Hauptevaluation werden die Antworten zunächst auf ihre Korrektheit geprüft. Hierzu werden alle Antwortbögen manuell erfasst und ausgewertet. Im Anschluss daran wird als zweites Kriterium die Bearbeitungszeit der korrekten Antworten erfasst und der Mittelwert daraus jedem Task und jedem Graph zugeordnet. Falsch abgegebene Antworten gehen nicht in die Berechnung mit ein, um eine Verfälschung der Ergebnisse zu vermeiden. Allerdings werden falsche Lösungen, die aufgrund des Zeitlimits entstanden sind, gesondert gefasst und in der Auswertung als solche gekennzeichnet. Nach der Analyse der Korrektheit und der jeweiligen Durchschnittszeit werden noch einige taskspezifische Sonderauswertungen vorgenommen, die zu interessanten Erkenntnissen führen.

3.3.1 Auswertung der Korrektheit der Antworten

Bei der Analyse der Korrektheit oder Qualität der gegebenen Antworten betrachten wir zunächst die Ergebnisse für jeden einzelnen Task separat. Hierbei ist zu beachten, dass Task 1 in dieser Auswertung nicht erscheint, da hier die Ergebnisse aufgrund des geringen Schwierigkeitsgrades nicht für eine Differenzierung zwischen richtigen und falschen Lösungen geeignet sind. In der gesamten Evaluation wurde Task 1 bei 324 Einzelfragen nur dreimal falsch beantwortet, was einer Quote von unter 1% entspricht. Daher sind die falschen Antworten hier als eine kurzzeitige Unkonzentriertheit der Probanden zu verstehen, die keine statistisch belegbaren Rückschlüsse auf mögliche Schwächen eines Layouts zulässt. In der Analyse der Durchschnittszeit und in der qualitativen Analyse im folgenden Unterkapitel wird Task 1 jedoch einige Erkenntnisse liefern.

Betrachtet man die Erfolgsquoten für Task 2 (siehe Abbildung 3.10), so werden die Ergebnisse der Vorstudie größtenteils bestätigt. Das kräftebasierte Verfahren liefert hier den besten Wert aller Layoutverfahren mit einer Quote von 93% korrekter Antworten. Dies erklärt sich zum einen daraus, dass das kräftebasierte Modell auf physikalischen Naturerscheinungen beruht, die dem Menschen im täglichen Leben begegnen. Viel bedeutender jedoch ist die Tatsache, dass ein Graph, der in diesem Design layoutet ist, am ehesten einer Landkarte entspricht, wie sie zum Beispiel Fluggesellschaften nutzen, um Flugverbindungen symbolisch darzustellen. Dadurch ist der menschliche Betrachter im Kontext eines Wegproblems hier am wenigsten irritiert. Begünstigend kommt zu diesem Umstand noch dazu, dass Kantenkreuzungen zu weit weniger Irritationen führen, als es bei einem Layout mit geschwungenen Kanten (Bögen) der Fall ist. Die Kanten machen hier keine Bögen oder Knicke, daher kann die Kante auch im Falle einer Kreuzung leichter weiterverfolgt werden. Das sehr gute Ergebnis der Vorstudie konnte allerdings auch das orthogonale Layoutverfahren bestätigen. Während in der Vorstudie die Erfolgsquote hier noch bei 100% lag, erreicht das Verfahren in der repräsentativen Studie immerhin noch einen Wert von 91% korrekter Antworten, wobei genau wie beim kräftebasierten Layout nur 7% der Antworten falsch sind und zusätzlich aber 2% keine Antwort im Zeitlimit abgeben konnten. Ob daraus eine generelle Tendenz herzuleiten ist, dass das orthogonale Verfahren gegenüber dem kräftebasierten zu einer Verlangsamung der Bearbeitungszeit führt, wird die Analyse der durchschnittlichen Bearbeitungszeiten ergeben. Der Erfolg dieses Layouts bei der Betrachtung des Wegproblems lässt sich erklären durch die besondere Übersichtlichkeit, die vom Algorithmus als oberste Vorgabe verwirklicht wird: Die Einbettung der Knoten erfolgt in einem Raster, auch die Kanten verlaufen entlang dieses Rasters und eine der Hauptdirektiven des

Algorithmus ist das Vermeiden von Kantenkreuzungen. Durch die Minimierung der Kantenkreuzungen nimmt die Zahl der Fehlerquellen stark ab. Ferner entspricht das orthogonale Layout ebenfalls stark schematischen Karten oder Plänen, wie wir sie aus dem täglichen Leben kennen. Ein Fahrplan für öffentliche Verkehrsmittel wie U- oder S-Bahn ist zum Beispiel oftmals stark angelehnt an das orthogonale Layout. Die Ergebnisse des hierarchischen Verfahrens können dagegen als Misserfolg dieses Layouts im Hinblick auf Task 2 betrachtet werden. Das Ergebnis von 59% korrekt abgegebener Antworten ist zwar eine deutliche Verbesserung gegenüber den Ergebnissen der Vorstudie, allerdings bildet dieses Verfahren im Vergleich zu den anderen getesteten Layouts eindeutig das Schlusslicht mit einem deutlichen Abfall der Quote. Nun muss zuerst der starke Unterschied zur Vorstudie erläutert werden, denn während die beiden anderen Verfahren ihre Ergebnisse in etwa bestätigen konnten, so hat das schichtbasierte Layout den Anteil der korrekten Antworten fast verdoppeln können. Hier zeigt sich, dass die Verbesserungen, die aus der Auswertung der Vorstudie hervorgingen, nämlich eine Erhöhung des vertikalen Abstandes der einzelnen Hierarchieschichten und damit eine Entzerrung des Layouts, eine große Wirkung zeigten. Eine Konsequenz für das Design hierarchischer Algorithmen im Hinblick auf die Zugänglichkeit ist damit, den Hierarchieschichten einen Mindestabstand zuzuordnen, der dazu führt, dass die Gesamtausdehnung des Graphen in vertikaler Richtung in etwa der in horizontaler Richtung entspricht. Auf diese Weise kann ein Kompromiss zwischen zu langen Kanten und zu waagrecht verlaufenden und spitze Schnittwinkel provozierenden Kanten erzielt werden, der die auftretenden Irritationen mindert. Aber auch diese Verbesserung des Layouts kann nicht darüber hinwegtäuschen, dass das hierarchische Verfahren sich für die Lösung eines Wegproblems weniger eignet. Die Schwierigkeiten für den menschlichen Betrachter liegen insbesondere in den Biegungen der Kanten und in den häufigen Kantenkreuzungen begründet. Auch der vergleichsweise hohe Wert von 6% der Antworten, die nicht innerhalb des Zeitlimits abgegeben werden konnten, spricht für eine mangelhafte Zugänglichkeit. Beachtet man hierbei noch, dass mit steigender Komplexität des Graphen die Ergebnisse sich überproportional stark verschlechterten (die Quote korrekter Antworten beim 20-Knoten Graph liegt gerade noch bei 5,5%), so muss man zu dem Schluss kommen, dass die hierarchische Darstellungsweise für beliebige Graphen ab einer Komplexität von ca. 15 Knoten ungeeignet ist, wenn damit Wegproblem - ähnliche Aufgaben bewältigt werden sollen.

Abbildung 3.11 zeigt die Gegenüberstellung der Erfolgsquoten der drei Layoutverfahren im Bezug auf Task 3. Trotz des als hoch eingeschätzten Schwierigkeitsgrades dieses Tasks kann das kräftebasierte Layout eine Erfolgsquote von 81% korrekter Antworten verbuchen. Hierbei kommt dem Verfahren zugute, dass es Knoten, die adjazent zueinander sind, in aller Regel in einer räumlichen Nähe darstellt (durch

den spring-embedder - Ansatz). In Task 2 ist genau diese Eigenschaft essentieller Bestandteil der Aufgabe, denn das Finden eines Teilgraphen entspricht dem Suchen nach äquivalenten Adjazenzbeziehungen im vorgegebenen Mustergraph und im eigentlichen Graph. Da nämlich eine ähnliche Form von Mustergraph und Teilgraph nicht gewährleistet werden kann, sind nur die Grapheigenschaften relevant. Auf diesen Umstand wurden alle Probanden im Vorhinein auch explizit hingewiesen. Aber gerade hier zeigt sich ein weiterer Vorteil des kräftebasierten Verfahrens, es erhält die „Form“ des Mustergraphen, sofern dieser einfach genug ist. Das heißt den Probanden war es hier möglich, nicht nur nach den Eigenschaften zu suchen, sondern zusätzlich nach einer Struktur, die dem kurz vorher gesehenen Bild entspricht. Durch die Kombination dieser Vorteile ist das gute Ergebnis zu erklären und die Zugänglichkeit des kräftebasierten Verfahrens im Bezug auf das Erkennen von Teilstrukturen belegt. Ein sehr schlechtes Ergebnis muss dagegen das orthogonale Verfahren verbuchen. Eine Erfolgsquote von nur 53% korrekt angegebener Antworten deutet hier darauf hin, dass das Verfahren ungeeignet dazu ist, bestimmte Teilstrukturen in einem komplexeren Graph zu erkennen. Diese These bestätigt auch die mit 35% höchste Fehlerquote aller Layouts. Wo liegen hier die Gründe? Betrachtet man die Aspekte, die zum guten Ergebnis des kräftebasierten Layouts beigetragen haben, so muss man feststellen, dass diese beim orthogonalen Layout ins Gegenteil verkehrt sind. Die oberste Prämisse des orthogonalen Layoutverfahrens ist die planare Einbettung des Graphen in das orthogonale Raster, wobei die Zahl der Kantenüberschneidungen minimal werden soll. Um dieses Ziel zu erreichen, ist eine vollständige Umordnung der Knoten notwendig. Dabei wird nicht auf eine enge Darstellung adjazenter Beziehungen geachtet, was dazu führt, dass zum Teil sehr lange Kanten entstehen. Eine adjazente Beziehung zweier Knoten ist jedoch sehr schwer ersichtlich, wenn die entsprechende Kante sehr lang ist. Dadurch liegt beim orthogonalen Layout eine Schwäche im Bezug auf Fragestellungen vor, bei denen direkte adjazente Beziehungen mehrerer Knoten von großer Wichtigkeit sind. Es wird interessant sein zu sehen, ob sich bei der Analyse von Task 4, der ähnliche Voraussetzungen hat, dieses Phänomen bestätigt. Betrachtet man die Werte des hierarchischen Layouts, so kann man feststellen, dass es mit 58% korrekt abgegebener Antworten etwas besser abschneidet als das orthogonale Layout. Hier fiel der Unterschied (+5%) zur Vorstudie weit weniger stark aus als bei Task 2. Dies ist damit zu erklären, dass das Entzerren des Layouts zwar zur besseren Übersichtlichkeit beiträgt, was den positiven Trend erklärt, aber auf der anderen Seite rücken benachbarte Knoten durch die Aufteilung in verschiedene Schichten weiter auseinander, was längere Kanten zur Folge hat. Diese Veränderung wirkt in die entgegengesetzte Richtung, fällt allerdings weit schwächer aus als beim orthogonalen Layout, da das hierarchische Verfahren trotz allem eine sehr kompakte Darstellung eines Graphen liefert und die Ausdehnung in vertikaler Richtung weniger stark ist

als die Folgen der (möglichst) planaren Einbettung beim orthogonalen Verfahren. Interessanter als die absolute Zahl falscher Antworten ist der Anteil an nicht beantworteten Aufgaben. Hier weist das hierarchische Layout mit Abstand den größten Wert auf. Dies ist ein weiteres Indiz dafür, dass das hierarchische Layout für beliebige Graphen generell schwerer zugänglich ist, denn insbesondere bei komplexer werdenden Graphen fällt, wie bei Task 2, die Rate der korrekten Antworten übermäßig stark (auf knapp 17% beim 20-Knoten Graph). Durch die starke Zunahme an gebogenen Kanten wird es für den Probanden schwerer, in einer fest vorgegebenen Zeit zu einem korrekten Ergebnis zu gelangen.

Eine Korrektur der Ergebnisse der Vorstudie fand allgemein bei Task 4 statt. Obwohl dieser von der Charakteristik her ähnlich aufgebaut ist wie Task 3, kann kein Layout die Ergebnisse der Vorstudie bestätigen (vergleiche Abbildung 3.12), alle drei Verfahren schneiden schlechter ab als in der Vorstudie. Hier scheint in der Vorstudie ein großer Lerneffekt für die recht guten Ergebnisse dieses Tasks verantwortlich zu sein, denn wie bereits erwähnt wurde Task 4 immer unmittelbar nach Task 3 bearbeitet. Da also in der Vorfrage bereits eine ähnliche Aufgabe am gleichen Graph zu bearbeiten war, wirkte Task 4 wie eine Wiederholungsfrage, die gerade geübte Analysemethoden konnte erneut (größtenteils erfolgreich) angewendet werden, zumal auch die Knotennamen die gleichen waren wie vorher. In der Hauptevaluation zeigen die Maßnahmen zur Vermeidung eines Lerneffektes (Permutation von Knotennamen, der Reihenfolge der Fragen und der Reihenfolge der Frageblöcke) eine deutliche Wirkung. Beim kräftebasierten Layout fällt die Quote korrekt beantworteter Tasks von 87% in der Vorstudie auf 76%. Damit ist dieses Verfahren allerdings immer noch deutlicher Spitzenreiter bei diesem Task. Die Erklärung hierfür ist ähnlich gelagert wie bei Task 3: Durch die enge Darstellung adjazenter Knoten ist es besser möglich, zusammenhängende Knotenstrukturen mit wenigen Blicken zu erkennen. Insbesondere Cliques werden in einem kräftebasierten Layout meist in einer sehr charakteristischen Art angeordnet, bei der hier zu findenden Viererclique ist dies eine rautenförmige Struktur mit einem Kantenkreuz in der Mitte. Es ist also möglich, zunächst nach einer ähnlichen Struktur zu suchen und hat man einen Kandidaten gefunden, so prüft man die Cliques - Eigenschaften der entsprechenden Knoten. Der Nachteil dieses Verfahrens wird deutlich, wenn man die falschen Antworten näher betrachtet: Hat man eine Gruppe von Knoten als mögliche Clique anvisiert, so neigt man dazu, bestimmte störende Knoten auszublenden. Hier wurde oftmals eine Lösung genannt, bei der noch ein Knoten in der Mitte der „Raute“ vorhanden ist, der allerdings einen Zwischenknoten des „Kantenkreuzes“ in der Mitte der Raute darstellt. Entweder wurde dieser Knoten fälschlicherweise ignoriert, oder es wurde eine Lösung mit fünf Knoten angegeben, die somit ebenfalls nicht korrekt ist. Das orthogonale Ergebnis bricht noch stärker ein, als es bei Task 3 der Fall ist. Wie ist

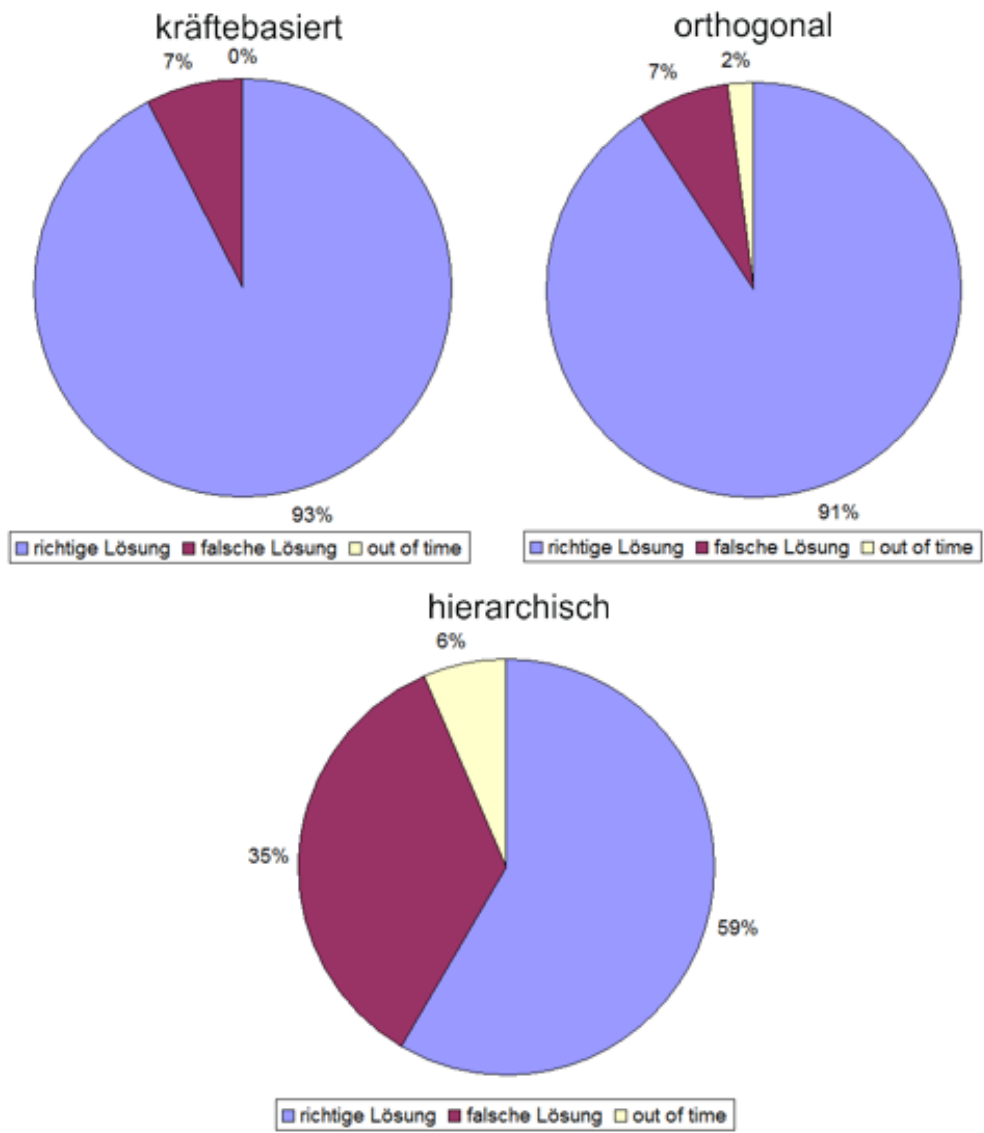


Abbildung 3.10: Erfolgsquoten Task 2

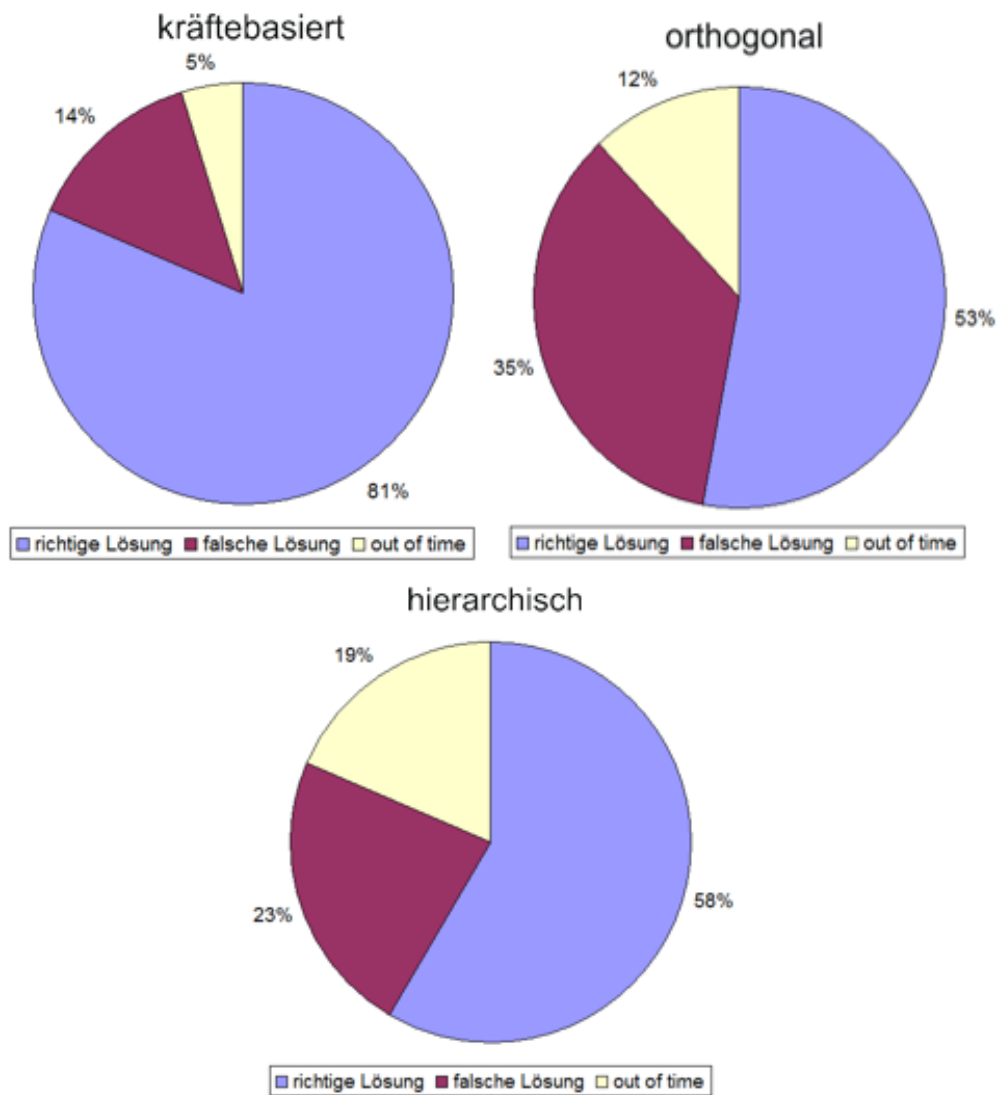


Abbildung 3.11: Erfolgsquoten Task 3

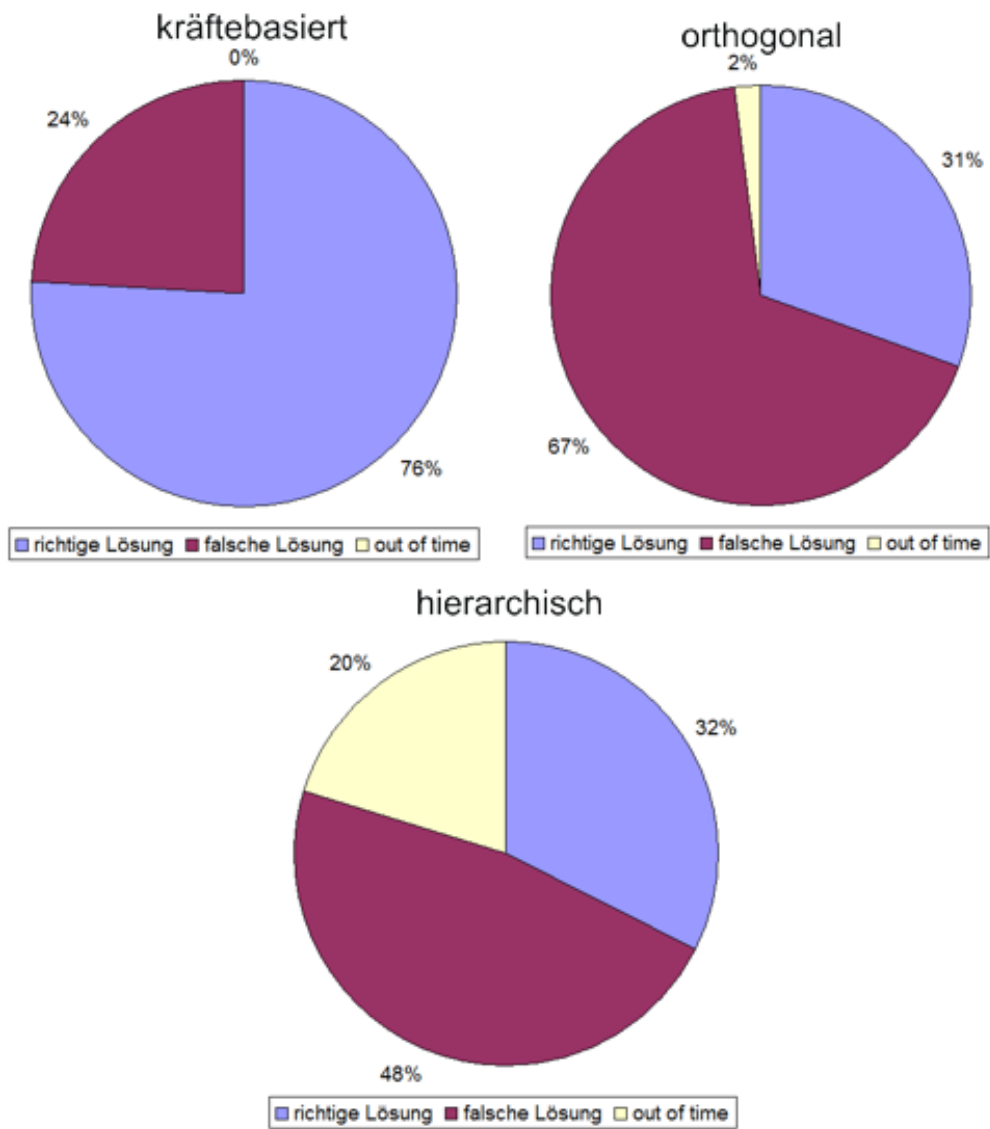


Abbildung 3.12: Erfolgsquoten Task 4

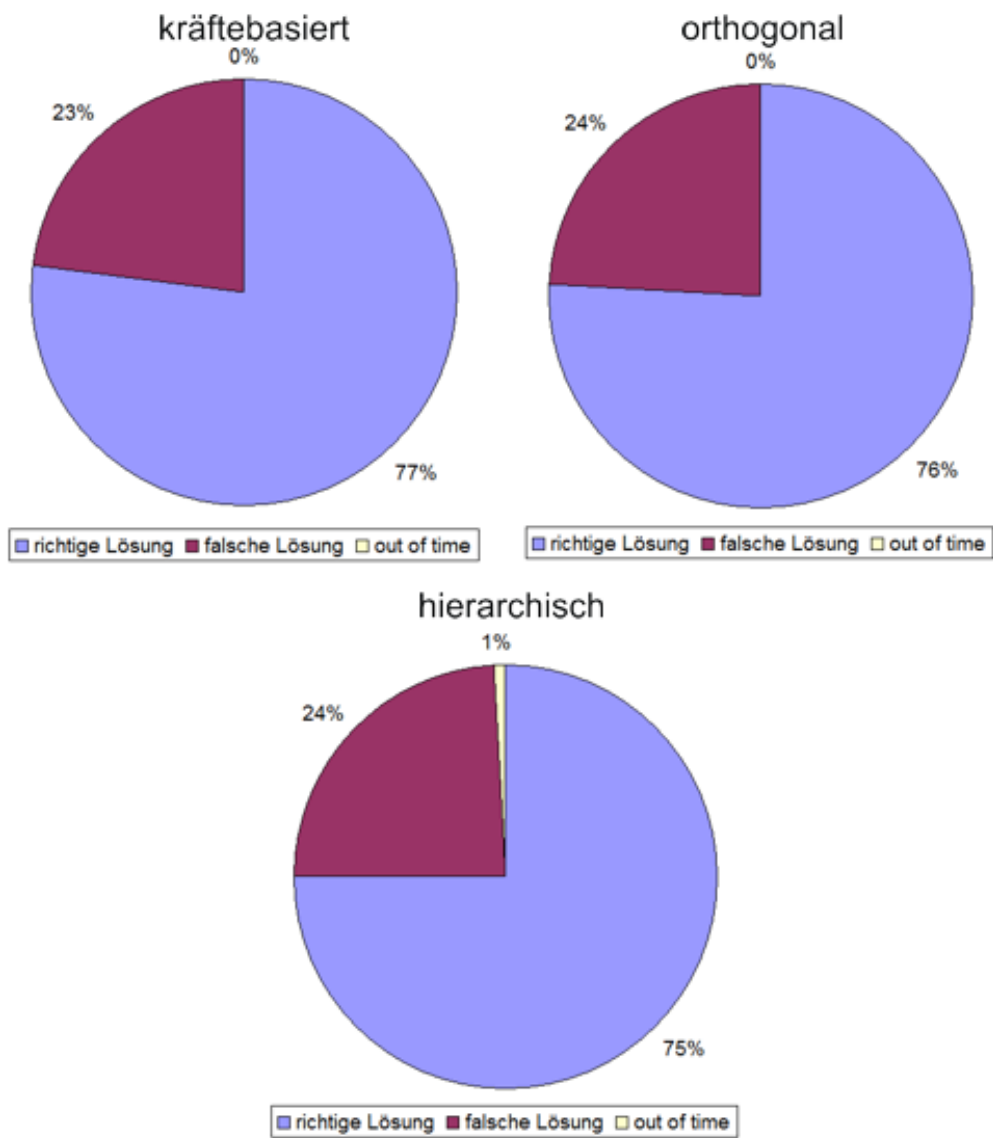


Abbildung 3.13: Erfolgsquoten Task 5

dies zu erklären? Hierzu fällt zunächst auf, dass 67% der Fragen falsch beantwortet sind. Es ist also keineswegs so, dass die Probanden zu lange überlegt bzw. gesucht hätten (der Anteil der nicht beantworteten Fragen liegt lediglich bei 2%), sondern es wurden falsche Lösungen angegeben. Scheinbar führt das orthogonale Layout, nach Eliminierung des Übungseffektes, dazu, dass Knoten als adjazent erscheinen, die es allerdings nicht sind. Wohingegen echte Vierercliquen nicht erkannt werden, da die betreffenden Knoten nicht als adjazent erkenntlich sind. Die Gründe liegen auch hier ähnlich gelagert wie bei Task 3. Durch die (versuchte) planare Einbettung in das orthogonale Gitter, verbunden mit der Kreuzungsminimierung, entstehen sehr lange Kanten, so dass eine Übersicht über adjazente Knoten nur schwer möglich ist. Task 4 basiert allerdings noch stärker als Task 3 auf den adjazenten Beziehungen zwischen Knoten, was dazu führt, dass dieses Layout unbrauchbar für diesen Aufgabentyp ist. Ein ähnlich schlechtes Ergebnis liefert auch das hierarchische Layout und dennoch bringt eine genaue Betrachtung Unterschiede zum orthogonalen Verfahren zutage. Denn während die Quote der korrekten Antworten mit 32% nur unwesentlich besser liegt als beim orthogonalen Layout, so ist der Anteil der aufgrund von Zeitmangel nicht beantworteten Fragen hier signifikant hoch und bestätigt damit die Thesen von Task 3. Durch gebogene Kanten, eine Vielzahl von Kreuzungen und die Verteilung adjazenter Knoten auf die unterschiedlichen Hierarchieebenen verliert das Layout bei starker Adjazenz-Bezogenheit der Aufgabe an Übersichtlichkeit. Besonders der dritte Punkt ist hier wichtig, denn aufgrund der Vorgabe, dass keine zwei adjazenten Knoten auf einer Hierarchieebene angeordnet sein dürfen, resultiert daraus eine Verteilung der Knoten einer Viererclique auf vier verschiedene Ebenen und damit eine vertikale Verzerrung.

Eine weitere teilweise Veränderung der Ergebnisse zwischen Vorstudie und Hauptstudie liefert die Analyse von Task 5 (Abbildung 3.13). Schnitt hier in der Vorstudie das kräftebasierte Layout noch deutlich schlechter ab als die beiden anderen Verfahren, so liegt es hier gleichauf. Damit leistet sich dieses Verfahren in keiner Disziplin eine echte Schwäche, es ist von der Quote der korrekten Antworten her immer mindestens so gut wie die konkurrierenden Layouts. Eine Erklärung für diese Veränderung kann nicht gegeben werden und müsste wohl eher in der Vorstudie gesucht werden, denn das Ergebnis der Hauptstudie entspricht dem ex ante erwarteten Resultat, wohingegen die überraschende These der Vorstudie widerlegt werden konnte. Das orthogonale und das hierarchische Layout können ihre Quoten aus der Vorstudie hingegen bestätigen. Der interessanteste Punkt an dieser Analyse von Task 5 ist allerdings die Tatsache, dass für die Feststellung des maximalen Grads das verwendete Layoutverfahren keine Bedeutung hat, zumindest was den Anteil an korrekten Antworten angeht. Ob die Bearbeitungszeiten variieren, werden wir später sehen. Dieser Umstand lässt sich nur daraus erklären, dass die Probanden

bei der Lösung dieser Aufgabe die Kanten (obwohl diese natürlich essentieller Teil dieses Tasks sind) weitestgehend ausblenden, einzig die unmittelbare Umgebung der Knoten, also das erste Stück austretende Kante, ist für die Lösung relevant. Die Darstellung adjazenter Beziehungen, die sich bei Task 3 und 4 als ein Gütekriterium für die Qualität eines Layouts erwiesen hat, verliert bei diesem Task, der jeden Knoten nur isoliert betrachtet, jegliche Bedeutung. Hierdurch wird deutlich, dass die Qualität eines Layouts immer Task - bezogen betrachtet werden muss, da unterschiedliche Tasks auch unterschiedliche Herangehens- und Betrachtungsweisen mit sich ziehen. Ob sich diese Tendenzen weiter verfestigen, wird die Analyse der Bearbeitungszeiten zeigen.

3.3.2 Auswertung der Bearbeitungszeiten

Die Übersicht über die durchschnittlichen Antwortzeiten soll zunächst ebenfalls taskspezifisch erfolgen. Abbildung 3.14 liefert hierzu die Übersicht zu Task 1. Bei diesem Task ist darauf zu achten, dass die Skalierung der Durchschnittszeitenachse hier den durchgängig sehr kurzen Antwortzeiten angepasst wurde, damit die Abbildung verwertbar wird. Bei allen anderen Übersichten ist die Skalierung untereinander identisch. Abbildung 3.14 zeigt zunächst, dass die durchschnittliche Bearbeitungszeit layoutunabhängig ist. Denn sowohl bei den 10-Knoten Graphen als auch bei den 15-Knoten Graphen liegen gleich komplexe Abbildungen jeweils gleich auf. Auch bei den 20-Knoten Graphen sind die durchschnittlichen Bearbeitungszeiten quasi identisch, der Ausreißer beim kräftebasierten Layout lässt sich dadurch erklären, dass der gesuchte Knoten vom Layoutalgorithmus am äußersten oberen Rand platziert wurde und daher nicht in den Bereichen zu finden ist, die normalerweise zuerst durchsucht werden, nämlich Bereiche mit einer größeren Knotendichte. Diese Layoutunabhängigkeit bestätigt die These, die gerade bei Task 5 aufgestellt wurde, nämlich eine Bedeutungslosigkeit des Layouts, solange nicht die Beziehungen zwischen Knoten betrachtet werden, sondern jeder Knoten ausschließlich isoliert betrachtet wird. Dies ist bei der Suche nach einem bestimmten Knoten natürlich der Fall. Betrachtet man die marginalen Unterschiede etwas genauer, lässt sich eine leichte Tendenz erkennen, dass das hierarchische Layout bei jeder Graphgruppe einen minimalen Vorteil hat, was durch die Kompaktheit dieses Layouts zu erklären sein dürfte. Allerdings müsste diese These anhand von komplexeren Graphen getestet werden, die Grundtendenz der Bedeutungslosigkeit des Layouts für diesen Task - Typ bleibt bestehen.

Wesentlich differenziertere Durchschnittszeiten erhält man bei der Betrachtung von Task 2 in Abbildung 3.15. Während die 10-Knoten Graphen von ihrer Komplexität her noch so einfach sind, dass bei allen Layouts die Erfassung des Weges in nahezu gleicher Zeit erfolgt, so zeigen sich bei den 15- und 20-Knoten Graphen deutlichere Tendenzen. Hier wird einmal mehr belegt, was sich auch in der Analyse der Erfolgsquoten von Task 2 bereits gezeigt hat: Das hierarchische Layout ist für einen solchen Aufgabentyp ungeeignet. Nicht nur die hohe Fehlerquote belegt dies, sondern auch die langen Bearbeitungszeiten. Bereits bei den 15-Knoten Graphen benötigten die Probanden trotz der geringen Komplexität im Durchschnitt drei Sekunden länger, um einen Weg zu bestimmen, als beim orthogonalen Verfahren. Das kräftebasierte Verfahren erweist sich um nochmal zwei Sekunden schneller. Beim 20-Knoten Graph wird deutlich, dass die Spanne mit steigender Komplexität noch größer wird. Während das kräftebasierte und das orthogonale Verfahren in etwa gleichauf liegen, benötigten die Probanden bei der hierarchischen Alternative im Durchschnitt 13 Sekunden länger. Zieht man noch die schlechte Quote der korrekten Antworten hinzu (deren Bearbeitungszeit nicht in die Durchschnittszeit eingerechnet ist, um die Ergebnisse nicht zu verfälschen), so wird deutlich, dass diese Art der Darstellung durch die gebogenen Kanten und die Vielzahl an Kreuzungen (mit zum Teil sehr spitzen Schnittwinkeln) bei einem beliebigen Graphen im Kontext des Wegproblems nicht an die Zugänglichkeit der anderen Verfahren heranreichen kann.

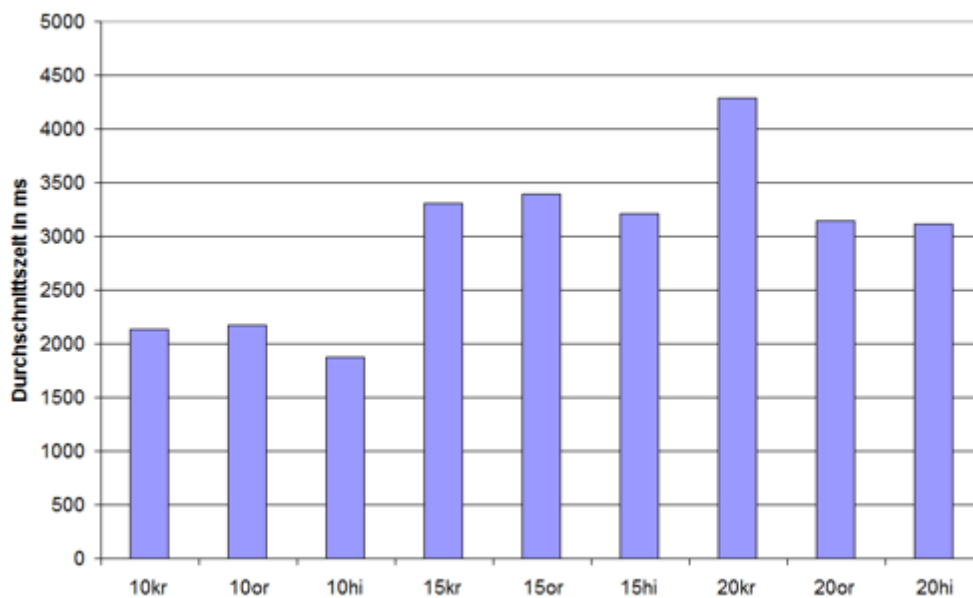


Abbildung 3.14: Durchschnittliche Bearbeitungszeit in Sek. (Task 1)

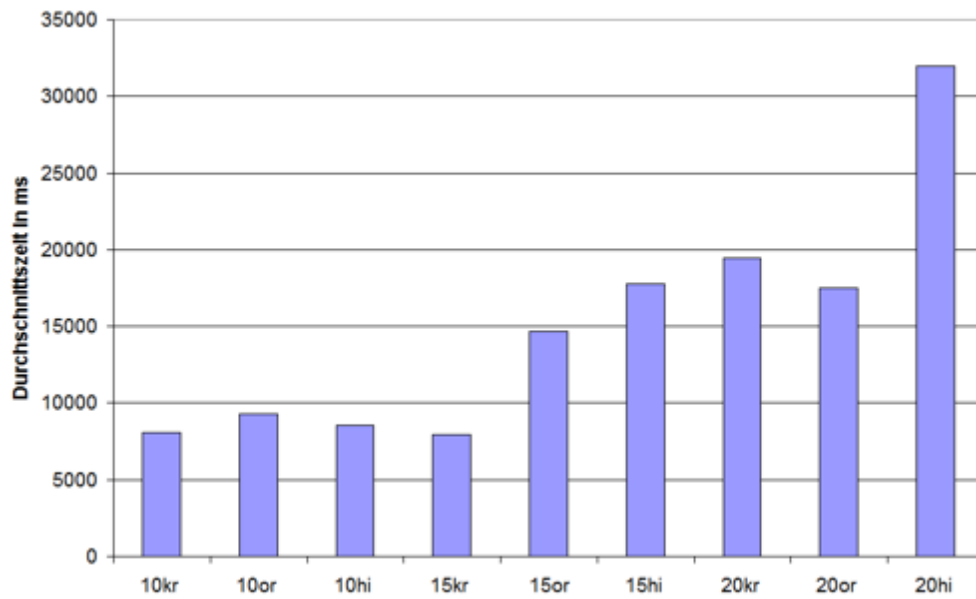


Abbildung 3.15: Durchschnittliche Bearbeitungszeit in Sek. (Task 2)

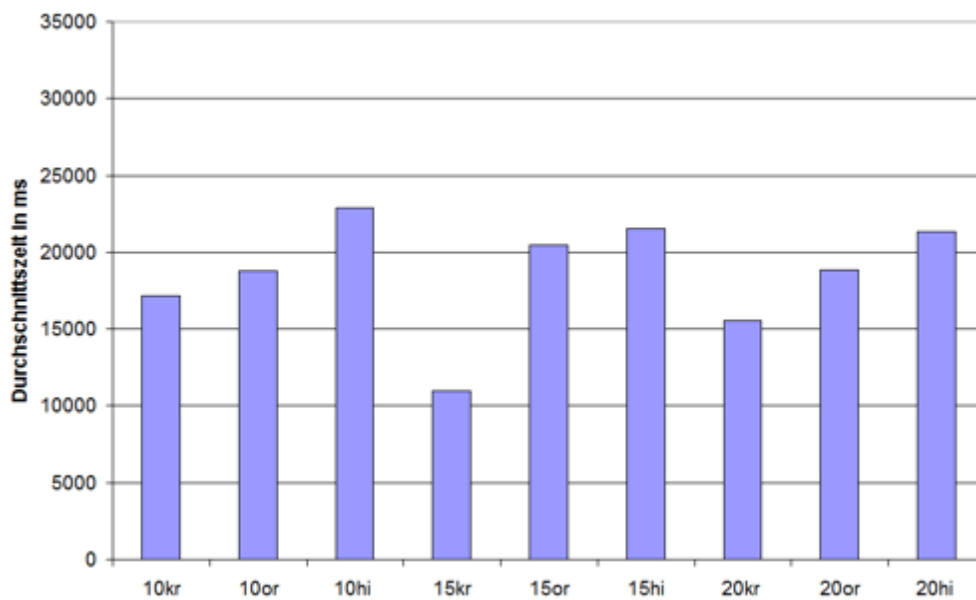


Abbildung 3.16: Durchschnittliche Bearbeitungszeit in Sek. (Task 3)

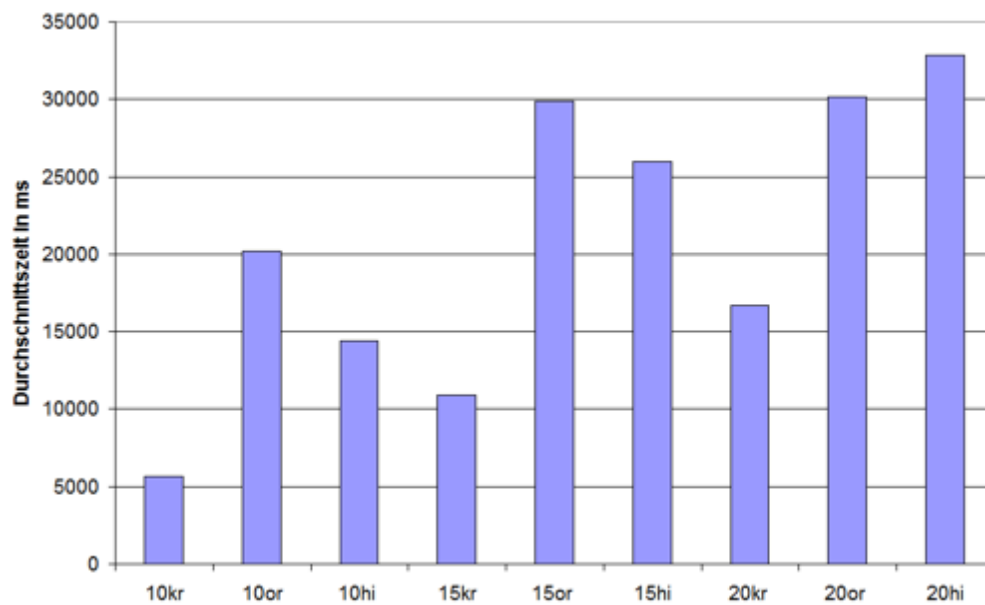


Abbildung 3.17: Durchschnittliche Bearbeitungszeit in Sek. (Task 4)

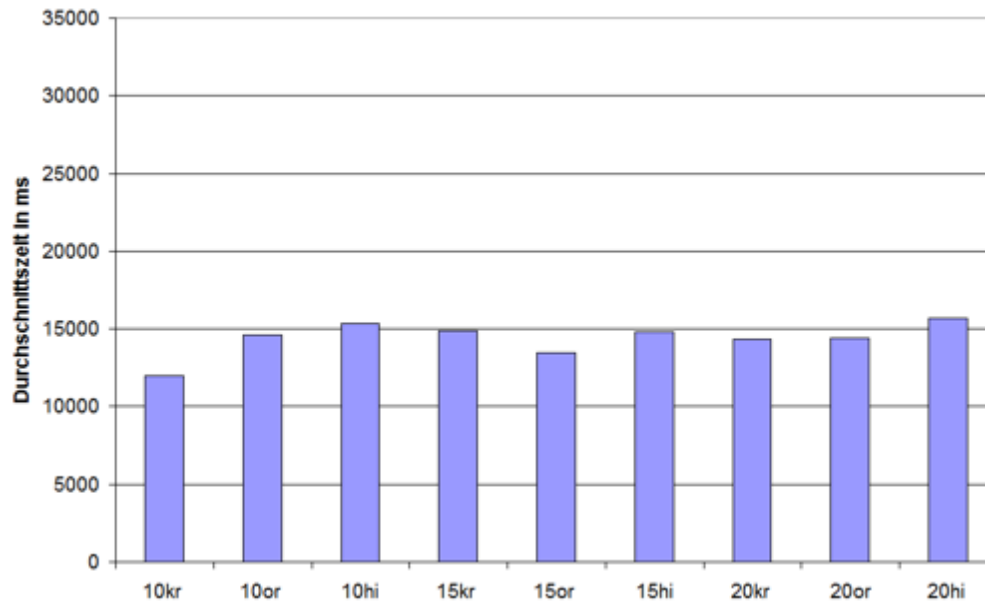


Abbildung 3.18: Durchschnittliche Bearbeitungszeit in Sek. (Task 5)

Abbildung 3.16 zeigt die Verteilung der durchschnittlichen Bearbeitungszeiten bei Task 3. Dabei ist zu erkennen, dass die kräftebasierte Darstellung bei allen Graphkomplexitäten die schnellste Zeit liefert. Dies entspricht den Thesen, die auf der Quote der korrekten Antworten beruhen. Denn in einem kräftebasierten Layout ist es wie bereits beschrieben eher möglich, nach einer bestimmten „Form“ oder Struktur zu suchen, da dieses Verfahren die Form besser bewahrt als die anderen Verfahren. Dadurch wird ein schnelles zweistufiges Suchverfahren ermöglicht, bei dem die Probanden zunächst eine Kandidatengruppe von Knoten suchen und anschließend nur noch diese Knoten überprüfen müssen. Bei den anderen Verfahren muss nach der rein syntaktischen Information gesucht werden, die durch den Vergleichsgraph vorgegeben ist, da diese die Struktur stark verändern. Eine weitere Auffälligkeit bei der Betrachtung von Abbildung 3.16 ist, dass die Bearbeitungszeiten offenbar unabhängig von der Komplexität des Graphen sind. Zwar muss hierbei beachtet werden, dass in die Durchschnittszeit der komplexeren Graphen weniger Einzelzeiten eingehen, da die Fehlerquote bei komplexeren Graphen weit höher liegt, allerdings kann die These vertreten werden, dass die Bearbeitungszeiten relativ unabhängig von der Graphgröße sind. Dies kann damit begründet werden, dass der Proband in kurzer Zeit einige Knotengruppen als Kandidaten auswählt (oder es beim orthogonalen und hierarchischen Layout zumindest versucht) und dann nur noch diese überprüft (und bei den beiden letztgenannten Layouts oftmals scheitert).

Dieses Scheitern tritt noch deutlicher zutage, wenn man den verwandten Task 4 in Abbildung 3.17 betrachtet. Hier erweist sich die Eigenschaft des kräftebasierten Verfahrens, eine Clique in einer sehr typischen Art und Weise abzubilden, als ein deutlich messbarer Vorteil in den Bearbeitungszeiten gegenüber den anderen Verfahren. Die Probanden benötigen für diese Aufgabe im Schnitt gerade halb so lange, wenn sie eine kräftebasierte Abbildung betrachten, als wenn sie einen anders layouteten Graph betrachten. Auch hier kann der Proband wiederum nach einer bestimmten Struktur im Graph suchen und erst dann die einzelnen Knoten untersuchen, ob sie die Cliqueneigenschaft erfüllen. Die Kandidatenauswahl ist dabei wegen der obigen Eigenschaft sehr erfolgversprechend. Das orthogonale Verfahren erweist sich wie bei der Analyse der Korrektheit bei den weniger komplexen Graphen als langsamstes Verfahren, wird mit steigender Komplexität aber in dieser Eigenschaft vom hierarchischen noch übertroffen. Dies ist damit zu begründen, dass bei wenig komplexen Graphen die Anzahl der Kanten noch gering ist, wodurch die Überprüfung der Kanten auf die Cliqueneigenschaft noch leichter möglich ist. Steigt allerdings die Komplexität und damit auch die Kantenzahl an, kommt es beim hierarchischen Verfahren zu einem starken Anstieg der Kantenkreuzungen, während beim orthogonalen Layout weiterhin die Vermeidung von Kreuzungen höchste Priorität genießt. Dadurch wird das orthogonale Verfahren bei größeren Graphen erfolgver-

sprechender sein als das hierarchische, allerdings dürfte der Abstand dieser beiden zum kräftebasierten Layout weiter anwachsen.

Bei dem Vergleich der Korrektheit der verschiedenen Layoutverfahren im Bezug auf Task 5 wurde die These formuliert, dass ein solcher Task, bei dem jeder einzelne Knoten isoliert betrachtet werden kann, weitestgehend layoutunabhängig ist, da die Layoutvarianten hauptsächlich die Darstellung von Beziehungen von Knoten untereinander beeinflussen. Betrachtet man Abbildung 3.18 kann man nur zu dem Schluss kommen, dass diese These auch durch die Analyse der durchschnittlichen Bearbeitungszeiten gestützt wird. Alle Einzelfragen weisen unabhängig vom betrachteten Layout in etwa die gleiche Durchschnittszeit auf. Dabei ist es um so erstaunlicher, dass dieser Task offenbar nicht nur vom Layout unabhängig ist, sondern auch von der Komplexität der Graphen, denn die 20-Knoten Graphen haben keine höhere Bearbeitungszeit als die 10-Knoten Graphen. Wie ist das zu erklären? Dadurch, dass die Probanden nicht alle Knoten betrachten, was eigentlich für diese Aufgabe nötig wäre, da ja ein globales Maximum angegeben werden soll, sondern sich gezielt einige wenige Knoten auswählen, deren Grad genauer bestimmt werden soll, spielt die Graphgröße nur eine untergeordnete Rolle. Die Anzahl der „Kandidaten“ steigt weit weniger schnell als die Komplexität des Graphen. Selbst eine Erhöhung des Durchschnittsgrads bedeutet nicht zwangsläufig, dass die Bearbeitungszeit steigt, da der Durchschnittsgrad ebenfalls eine globale Variable ist, so dass die Zahl der „Kandidaten“ nicht steigt, sondern nur der Grad (natürlich kann es bei einem ungleichmäßigen Hinzufügen von Kanten neue Kandidaten geben, aber dadurch werden andere Knoten als Kandidaten ausfallen).

Alternativ zur taskbezogenen Analyse der durchschnittlichen Bearbeitungszeiten kann man auch einen layoutbezogenen Vergleich in Erwägung ziehen. Dieser wird für das kräftebasierte Layoutverfahren in Abbildung 3.19 gezeigt. Hier wird mit steigender Komplexität der Graphen deutlich, wie sich die Bearbeitungszeiten bei den einzelnen Tasks verschieben. Während, wie schon gesehen, Task 1 und 5 unabhängig von der Komplexität des Graphen eine relativ konstante Bearbeitungszeit aufweisen können, da hier jeder Knoten isoliert betrachtet wird und somit kaum beeinflusst wird von der wachsenden Komplexität um ihn herum, verändern sich die Werte der anderen Tasks mit steigender Komplexität stetig. Sieht man von einem Ausreißer (Task 3 beim 10-Knoten Graph) ab, so steigt die Bearbeitungszeit der Tasks 2 bis 5 deutlich an, bei Task 2 ab einer Komplexität von 20 Knoten sogar übermäßig stark. Task 3 und 4 entwickeln sich aufgrund ihrer thematischen Verwandtschaft erwartungsgemäß parallel. Beim orthogonalen Verfahren in Abbildung 3.20 ist eine ähnliche Entwicklung festzustellen. Hier erweist sich Task 4 als der durchweg zeitaufwändigste Task. Die Gründe dafür wurden oben bereits ausgeführt, näm-

lich die schlechte Darstellung adjazenter Beziehungen zwischen den Knoten. Soll der orthogonale Algorithmus also für einen Graphen eingesetzt werden, bei dem ein Adjazenz-bezogener Kontext von großer Wichtigkeit ist, so muss zur Verbesserung der Zugänglichkeit eine Anpassung des Algorithmus vorgenommen werden, die darauf abzielt, in Beziehung stehende Knoten auch nahe beieinander abzubilden. Ähnliches gilt natürlich für Task 3, der bei jeder Graphkomplexität die zweitlängste Bearbeitungszeit vorweist. Task 5 weist anfangs einen recht hohen Wert auf, allerdings haben wir bereits festgestellt, dass dieser Wert auch bei steigender Komplexität konstant bleibt, die Gründe wurden erläutert. Daher wird Task 5 von Task 2 bereits ab einer Komplexität von 15 Knoten in der Bearbeitungszeit übertroffen. Die weitere Entwicklung ist hier zwar auch steigend, fällt aber weniger intensiv aus als dies zum Beispiel beim hierarchischen Layout der Fall ist. Dessen Gesamtentwicklung verdeutlicht Abbildung 3.21. Hier bestätigt sich ebenfalls die Einschätzung der Tasks 3 und 4 als Aufgaben mit dem höchsten Schwierigkeitsgrad. Diese Fragen benötigen bereits beim 10-Knoten Graph eine hohe Bearbeitungszeit, diese steigt mit der Komplexität des Graphen kontinuierlich an. Entscheidend ist allerdings hier der Vergleich mit Task 2, denn im Vergleich der Tasks wird deutlich, dass diese Aufgabe in der Bearbeitungszeit mit steigender Komplexität überdurchschnittlich stark ansteigt und den Abstand zu Task 4, der die schlechteste Bearbeitungszeit hat, ständig verringert. Daher ist anzunehmen, dass bereits bei einer leichten Erhöhung der Graphkomplexität Task 2 einer längeren Bearbeitungszeit bedarf als Task 4. Diese Entwicklung lässt sich daraus erklären, dass bei Task 4 immer noch (wenn auch abgeschwächer als beim kräftebasierten Layout) bestimmte Kandidatengruppen gesucht werden und diese dann nur überprüft werden. Bei Task 2 hingegen müssen je nach Weglänge sehr viele Knoten und Kanten betrachtet werden, die sich zumal über den ganzen Graph verteilen können. Damit erhöht sich bei steigender Komplexität die Zahl der Kantenkreuzungen und der langen Kanten übermäßig stark.

3.3.3 Weitere Auswertungen

Neben der Quote der korrekten Antworten und der durchschnittlichen Bearbeitungszeit bietet vor allem die intensivere Analyse von Task 2 einige weitere Erkenntnisse, die an dieser Stelle erläutert werden sollen. Zur Erinnerung: Die Aufgabenstellung von Task 2 verlangte von den Probanden nur, einen Weg von einem vorgegebenen Startknoten zu einem ebenfalls festgelegten Endknoten zu bestimmen. Die Frage suchte ganz bewusst nicht den kürzesten Weg zwischen den betreffenden Knoten, viel mehr sollte festgestellt werden, ob eines der getesteten Layouts das Konzept

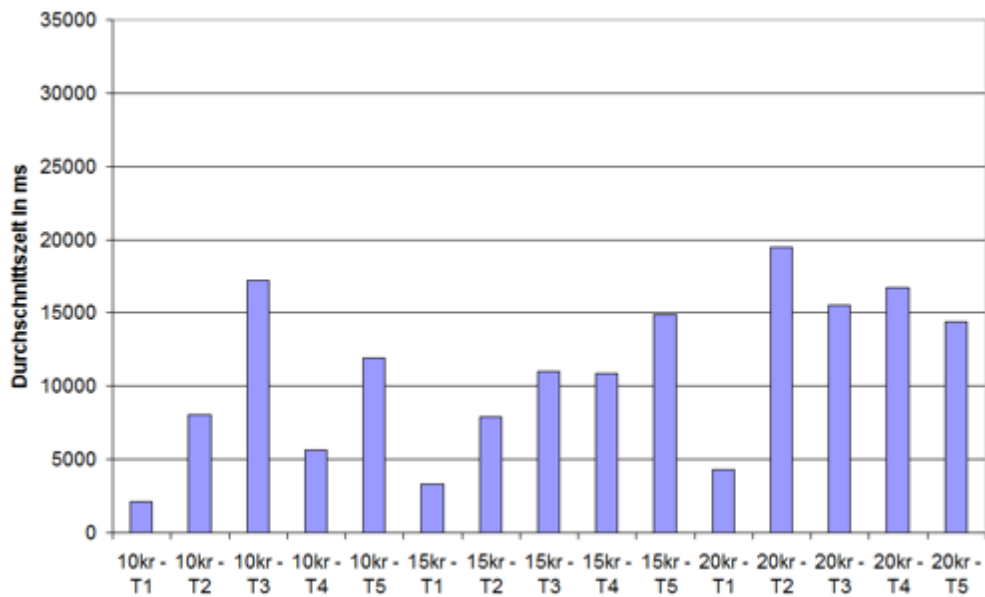


Abbildung 3.19: Durchschnittliche Bearbeitungszeit in Sek. (kräftebasiertes Layout)

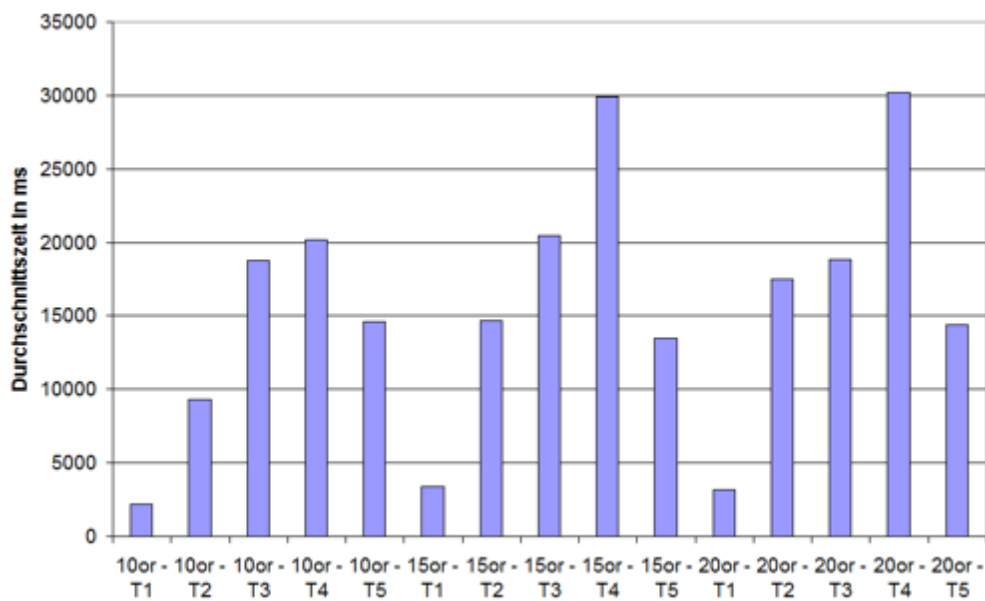


Abbildung 3.20: Durchschnittliche Bearbeitungszeit in Sek. (orthogonales Layout)

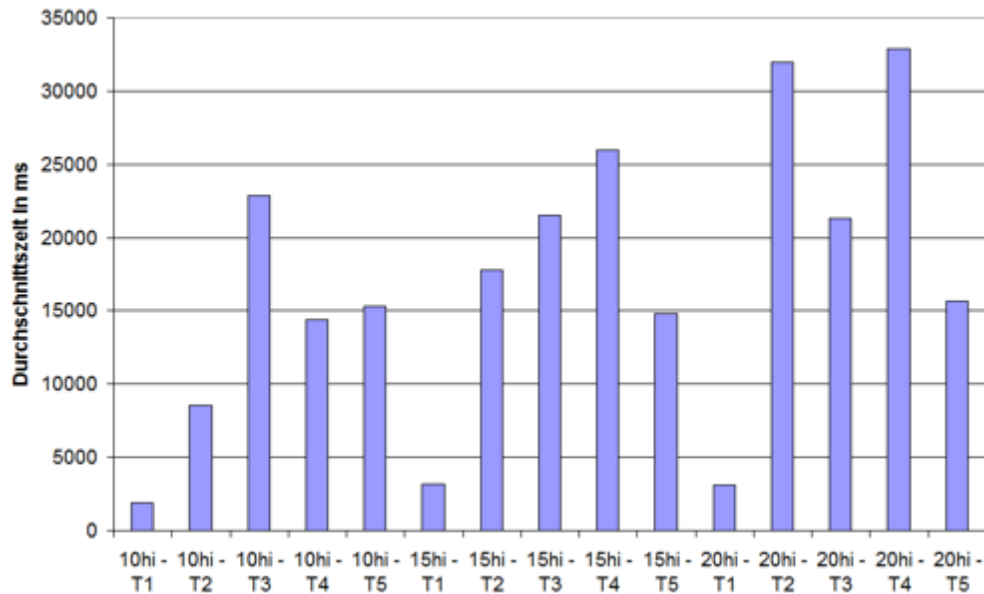


Abbildung 3.21: Durchschnittliche Bearbeitungszeit in Sek. (hierarchisches Layout)

„kürzester Weg“ implizit unterstützt. Um diese Frage zu beantworten, muss zunächst geklärt werden, was ein kürzester Weg im Rahmen dieser Evaluation ist. Der kürzeste Weg ist die Verbindung zwischen zwei Knoten, bei der die geringsten Pfadkosten anfallen. Da bei dieser Befragung ex ante keinerlei Aussagen über die Pfadkosten gemacht wurden, muss man davon ausgehen, dass uniforme Kantenkosten vorliegen. Man kann also o.B.d.A. davon ausgehen, dass die Kosten einer Kante „1“ betragen. Damit ist das Problem des kürzesten Pfades gleichzusetzen mit der Frage nach der Minimierung der Anzahl der Kanten des Pfades oder der Minimierung der Zwischenknoten, die Teil des Weges sind. Basierend auf dieser Annahme findet man in Abbildung 3.22 eine Übersicht über die Korrektheit der Antworten zu Task 2, wobei die richtigen Lösungen noch einmal unterteilt sind in „nur korrekte“ Antworten, in falsche Antworten und in korrekte Lösungen, die dem kürzesten Weg entsprechen (obwohl dieser nicht gefordert war).

Bereits beim ersten Betrachten von Abbildung 3.22 wird deutlich, dass bei allen drei Layoutverfahren ein völlig unterschiedliches Ergebnis zustande kommt. Wie bereits weiter vorne gesehen, erreicht das kräftebasierte Layout einen sehr guten Wert von 93% korrekter Antworten. Differenziert man nun bei diesen Antworten nach „nur korrekten“ Antworten und Lösungen mit der Nennung des kürzesten Weges, so zeigt sich eine weitere Stärke dieses Verfahrens: 73% der korrekten Ant-

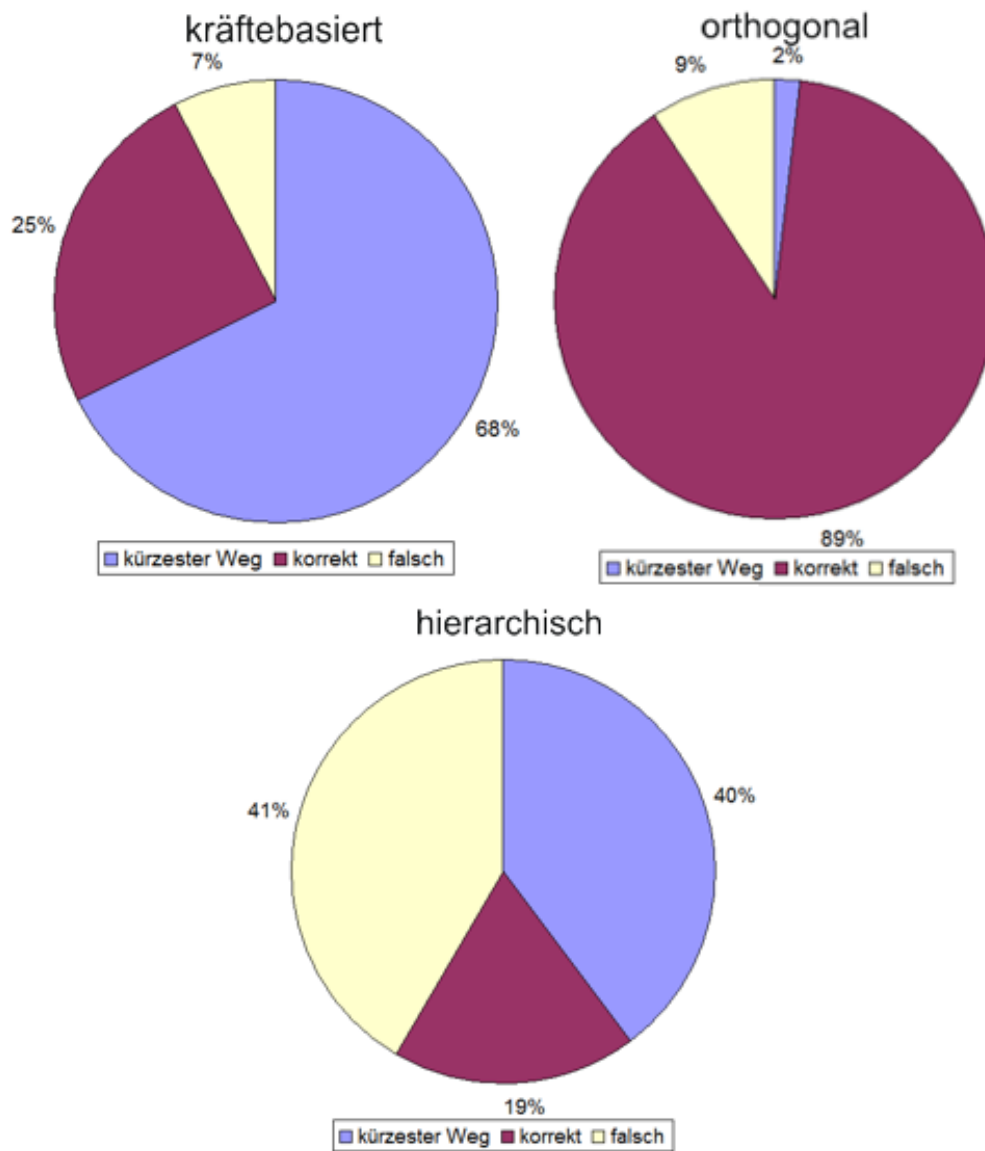


Abbildung 3.22: korrekte Antworten vs. kürzeste Wege (Task 2)

worten (bzw. 68% der gesamten Antworten) lieferten einen kürzesten Weg nach der obigen Definition. Da immer wieder darauf hingewiesen werden muss, dass die Probanden nicht dazu aufgefordert wurden, einen kürzesten Weg zu finden, kann man hier feststellen, dass das kräftebasierte Layout in hohem Maße das Finden eines solchen Weges durch einen menschlichen Betrachter unterstützt. Die Erklärung hierfür dürfte beim „spring-embedder“ zu suchen sein. Durch das zugrunde liegende physikalische Prinzip des Verfahrens wird gewährleistet, dass benachbarte Knoten in direkter Nähe zueinander abgebildet werden. Zudem kommen in einem solchen Graph ausschließlich gerade Kanten vor, wodurch das Erkennen eines direkten Weges für den menschlichen Betrachter wesentlich erleichtert wird. Darüber hinaus neigt das kräftebasierte Layout dazu, den Graph so abzubilden, dass der Proband einen (kürzesten) Weg verfolgen kann, der nahe an einer gedachten Geraden zwischen Start- und Zielknoten verläuft. In manchen Fällen, wie z.B. beim 20-Knoten Graph, sind Zwischenziele einzuplanen, aber eine grundsätzliche Übersichtlichkeit über die einzuschlagende Richtung bleibt stets erhalten.

Vergleicht man dieses Ergebnis mit der Darstellung des orthogonalen Layouts in Abbildung 3.22, so muss man deutliche Unterschiede feststellen. Betrachtet man nur die Quoten der korrekten Antworten, so erscheinen das kräftebasierte und das orthogonale Verfahren mit 93% zu 91% gleichauf. Erst eine Analyse des Anteils der kürzesten Wege an allen richtigen Antworten zeigt die großen Unterschiede zwischen den beiden Layoutvarianten. Die Probanden nannten nur in 2% der Fälle eine Antwort, die einem kürzesten Weg nach obiger Definition entspricht. Wie kann ein so großer Unterschied bestehen bei zwei Layoutverfahren, deren Gesamtquote an korrekten Antworten so dicht beieinander liegt? Die Antwort liegt einmal mehr im orthogonalen Algorithmus begründet, genauer in der Einbettung der Knoten im orthogonalen Gitter. Wie bereits bei der Analyse der Bearbeitungszeiten verdeutlicht, führt diese Einbettung, die vorrangig darauf optimiert ist Kantenkreuzungen zu vermeiden, dazu, dass benachbarte Knoten sehr weit voneinander weg abgebildet werden und dadurch durch sehr lange Kanten verbunden werden. Nun ist es für den menschlichen Betrachter ohne eine graphtheoretische Vorbildung nur schwer nachzuvollziehen, dass eine Kante mit der Länge von bis zu 20 cm die gleichen Kosten haben soll wie eine Kante, die gerade einmal 1 cm lang ist. Dies widerspricht völlig der täglichen Erfahrung, dass kurze Entfernungen geringe Kosten und weite Entfernungen hohe Kosten verursachen. Dass sich dadurch das orthogonale Layout nur bedingt eignet, wenn es tatsächlich um das Finden eines kürzesten Weges geht, wird dadurch deutlich. Gestaltet z.B. ein orthogonaler Algorithmus einen U-Bahn-Plan, kann es passieren, dass zwei Strecken, die zum gleichen Ziel führen, völlig unterschiedlich dargestellt werden: Eine Linie macht an jeder Zwischenstation einen Stop, während die andere Linie die beiden Punkte direkt verbindet. Wird diese

Direktverbindung durch den Algorithmus aber aufgrund von Platz- und Kantenoptimierungsgründen einmal komplett um den ganzen Plan herumgeführt, so wird sich jeder Fahrgast automatisch für die langsamere Strecke mit den Zwischenstationen entscheiden. Man kann also festhalten, dass sich das orthogonale Layout zwar gut dazu eignet, einen Weg von A nach B zu finden, allerdings erschwert es das Finden eines kürzesten Weges enorm durch die orthogonale Einbettung der Knoten, die ein anderes Optimierungsziel verfolgt.

Es wurde bereits mehrfach gezeigt, dass das hierarchische Layout eher als ungeeignet zur Lösung des Wegproblems angesehen werden muss. Doch wie sehr unterstützt dieses Layout implizit das Finden eines kürzesten Weges? Abbildung 3.22 liefert hier ein überraschendes Ergebnis. Denn obwohl nicht nach einem kürzesten Weg gefragt wurde, stellen 40% der Lösungen tatsächlich einen kürzesten Weg zwischen vorgegebenem Start- und Zielknoten dar. Rechnet man diesen Anteil um auf die Gesamtzahl der korrekt abgegebenen Lösungen, so erreicht das hierarchische Layout einen Wert von 68% Antworten, die einen kürzesten Weg beschreiben. Damit liegt das hierarchische Verfahren nur unwesentlich hinter dem guten Wert des kräftebasierten Verfahrens (73%). Daraus kann gefolgert werden, dass das hierarchische Layout es den Probanden zwar erschwert, überhaupt eine Lösung zu finden (insbesondere bei komplexeren Graphen), wenn sie allerdings doch eine Lösung gefunden haben, so ist die Wahrscheinlichkeit groß, dass es sich dabei um einen kürzesten Weg handelt. Das hierarchische Layoutverfahren unterstützt also in hohem Maße implizit das Finden eines kürzesten Weges, allerdings eignet es sich generell eher weniger für diesen Aufgabentyp, da die Übersichtlichkeit besonders bei komplexeren Graphen stark unter den gebogenen Kanten und der Vielzahl von Kreuzungen leidet. Betrachtet man das Ergebnis aller Layoutverfahren insgesamt, so ist es erstaunlich, wie unterschiedlich die implizite Unterstützung des Konzeptes „kürzester Weg“ ausfällt. Es wäre in einer künftigen Studie zu prüfen, ob bei dem expliziten Auftrag, einen kürzesten Weg zu finden, diese implizite Unterstützung sich auch in einer absoluten Erfolgsquote niederschlägt.

Zum Abschluss der quantitativen Analyse soll noch eine Übersicht über die Gesamterfolgsquoten der einzelnen Layoutverfahren gegeben werden. Hierzu dient Abbildung 3.23. Bei dieser Zusammenstellung werden alle Antworten der Probanden zu allen Tasks nach Graphlayouts differenziert betrachtet. Zusätzlich ist noch die Gesamterfolgsquote der Gesamtevaluation abgebildet. Die Aussagekraft dieser Abbildung im Hinblick auf spezielle Verbesserungen an den Layoutalgorithmen ist zwar gering, allerdings ermöglicht sie es, eine Tendenz festzustellen, welches Verfahren bei unbekanntem oder sehr breit gestreutem Task - Hintergrund am besten geeignet ist. Wie bereits erwartet, erweist sich das kräftebasierte Layout als der Allrounder

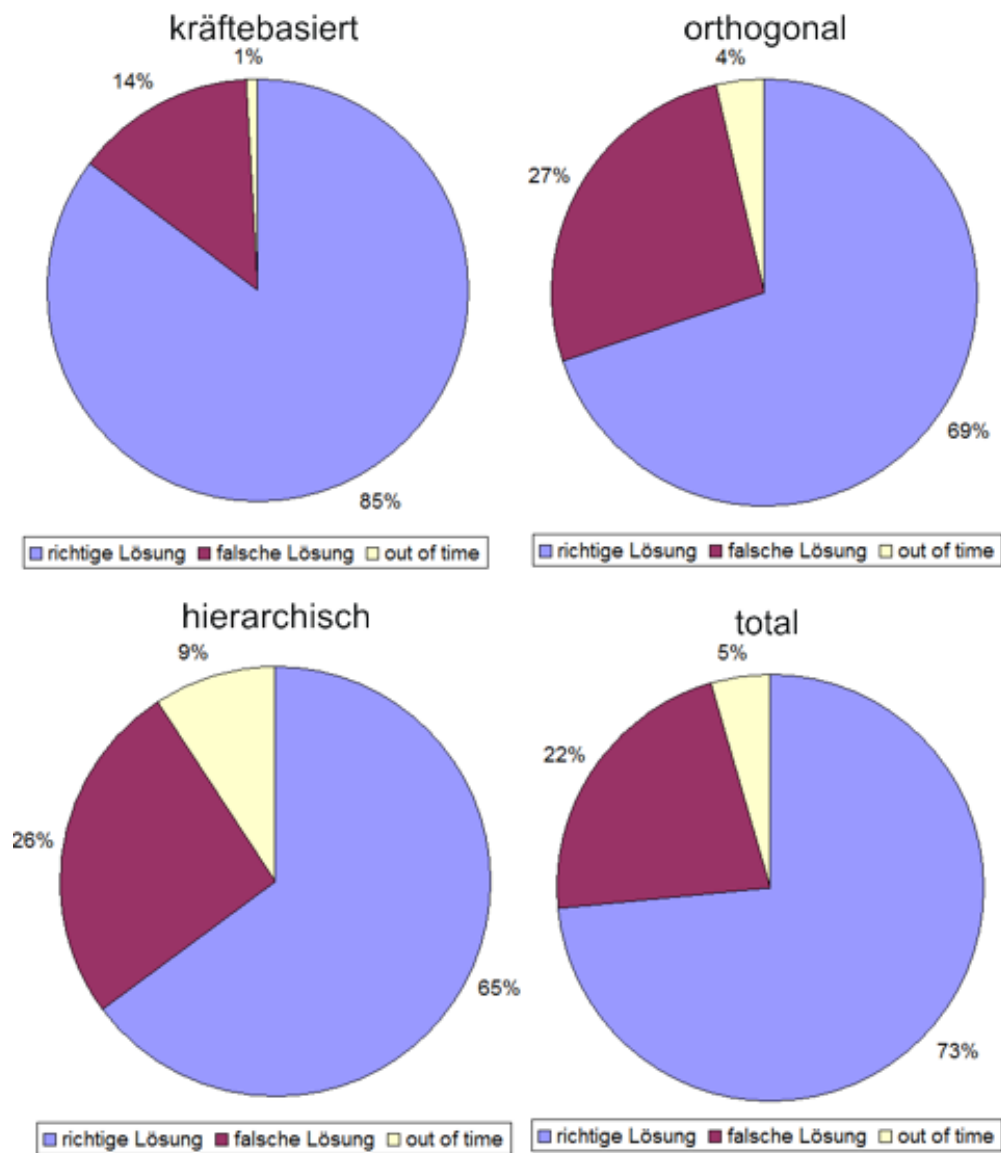


Abbildung 3.23: Gesamterfolgsquoten über alle Tasks

unter den Verfahren. Es zeigt in keiner Disziplin echte Schwächen, was zu einem überdurchschnittlich guten Ergebnis von 85% korrekt beantworteter Fragen führt. Der besonders niedrige Anteil an Fällen, bei denen innerhalb des Zeitlimits keine Antwort gegeben wurde, deutet zudem darauf hin, dass hier der Graph besonders schnell erfasst wird, so dass nur in Ausnahmefällen das Ausnutzen der vollen Zeitspanne nötig war. Deutlich dahinter findet man das orthogonale Verfahren mit einer Erfolgsquote von 69% richtiger Antworten. Im Vergleich zum kräftebasierten Verfahren fällt insbesondere auf, dass sich die Zahl der falschen Antworten fast verdoppelt hat. Auch die Zahl der aufgrund Zeitmangels nicht gegebenen Antworten hat sich auf 4% erhöht. Damit kann das orthogonale Verfahren im Bezug auf seine Taskunabhängigkeit nicht mit dem kräftebasierten Verfahren konkurrieren, allerdings haben wir auch gesehen, dass bei einzelnen isolierten Tasks die Werte durchaus vergleichbar sind. Das hierarchische Layout hingegen kann bei taskunabhängiger Betrachtung nicht als Allround - Verfahren empfohlen werden, insbesondere nicht, wenn beliebige Graphen layoutet werden. Während die Quote der falschen Antworten im Vergleich zum vorherigen Layout gleich bleibt, steigt die Quote der nicht abgegebenen Antworten auf den Rekordwert von 9%, was eindeutig darauf hinweist, dass (besonders bei komplexeren Graphen) die Zugänglichkeit für den menschlichen Betrachter stark zurückgeht, was nur durch einen hohen Zeitaufwand halbwegs kompensiert werden kann. Es ist durchaus vorzustellen, dass bei einer Beschränkung auf bestimmte Aufgaben und bei Verwendung von Graphen in einem speziellen semantischen Kontext, das hierarchische Layout seine Qualitäten ausspielen kann, diese Voraussetzungen müssen aber in jedem Fall gegeben sein, damit es sich gegenüber einem kräftebasierten oder auch orthogonalen Layout behaupten kann. Insgesamt gesehen hat sich das Design der Studie als Erfolg erwiesen, was sich auch in der Gesamtquote in Abbildung 3.23 zeigt. Mit 73% korrekter Antworten wurde ein sehr gutes Ergebnis erzielt, das darauf hindeutet, dass der Schwierigkeitsgrad und die zeitliche Beschränkung treffend gewählt waren und zu aussagekräftigen Ergebnissen führen. Während eine Tendenz korrekter Antworten Richtung 90% doch eher darauf hingedeutet hätte, dass der Schwierigkeitsgrad zu niedrig angesetzt war, hätte eine Verlagerung der Quote in Richtung der 50%-Marke auf zu schwierige Tasks hingedeutet, was wiederum die jeweilige Aussagekraft geschwächt hätte.

3.4 Qualitative Analyse

Neben der bisherigen Vorgehensweise zur Analyse der Versuchsreihe bietet die Verwendung des Eyetracking-Systems die Möglichkeit, einige ungeklärte Fragen oder

nicht hinreichend belegte Thesen durch eine direkte Analyse der Blickbewegungsdaten weiter zu klären. Wie bereits beschrieben, wurden pro Proband und pro Frageblock jeweils eine Aufzeichnung vorgenommen. Daraus ergibt sich eine Gesamtzahl von 108 einzelnen Aufzeichnungen, die es auszuwerten galt. Jede dieser Aufzeichnungen umfasst wiederum 15 einzelne Fragen, nach denen differenziert werden muss. Um aus dieser Vielzahl von Analysemöglichkeiten statistisch belegbare Erkenntnisse zu gewinnen, wurden bei der Auswertung pro Frage die Daten aller Probanden zusammengelegt. Das heißt, dass die Summe der Fixationen durch alle Probanden für jeden Punkt einer Abbildung berechnet und im Anschluss als eine *Heatmap* dargestellt wurde. Bei einer solchen Heatmap wird die Gesamtzahl der Fixationen durch einen Farbverlauf dargestellt, dabei verläuft die Farbskala von grün (wenige Fixationen) über gelb nach rot (sehr viele Fixationen). Dabei bietet CLEARVIEW die Möglichkeit, diese Skalierung anzupassen, um so sowohl globale als auch sehr fein abgestufte Auswertungen vornehmen zu können. Ferner wurden die Daten des Eyetracking-Systems dazu genutzt, bei falschen Antworten den Grund für den jeweiligen Fehler zu finden. Hierbei wurde bei häufig auftretenden Fehlern oder bei besonderen Schwierigkeiten bei einem bestimmten Layout oder einer bestimmten Frage die jeweilige Aufzeichnung des Probanden betrachtet, sowie die einzelnen Fixationen im betreffenden Zeitfenster analysiert. Zusätzlich lieferte die Analyse der Eyetrackerdaten im Bezug auf Task 5 einen Impuls für eine weitere, noch angehängte Zusatzstudie, die am Ende des Kapitels genauer erklärt und analysiert wird.

3.4.1 Heatmap-Analyse

Mit Hilfe der Heatmap-Analyse sollen einzelne Phänomene weiter analysiert werden, deren Existenz aufgrund der Ergebnisse der quantitativen Analyse zwar vermutet, aber noch nicht belegt werden kann. Eines dieser Phänomene ist ein so genannter „*Scheuklappeneffekt*“. Mit diesem Begriff soll das beobachtete Verhalten fast aller Probanden beschrieben werden, bestimmte (offensichtlich unwichtige (?)) Bereiche der Abbildung bei der Lösung der Aufgabe auszuklammern, ja völlig zu ignorieren. Ob diese Bereiche letztlich tatsächlich für die Lösung der Aufgabe unwichtig sind, zeigt die genauere Betrachtung der Heatmaps ebenfalls. Besonders gut geeignet zur Demonstration dieses Effektes sind die Auswertungen zu Task 2, da hier ein besonders zielgerichtetes Verhalten seitens der Probanden von Nöten war, um vom Start zum Zielknoten zu gelangen. Hier sind je nach Layout (nicht beachtenswerte) Umwege besonders leicht oder auch nur sehr schwer zu erkennen, wie sich bereits bei der quantitativen Analyse der impliziten Unterstützung des kürzesten Weges erkennen ließ. Da in dieser Arbeit schon aus Platzgründen nicht jede einzelne untersuchte

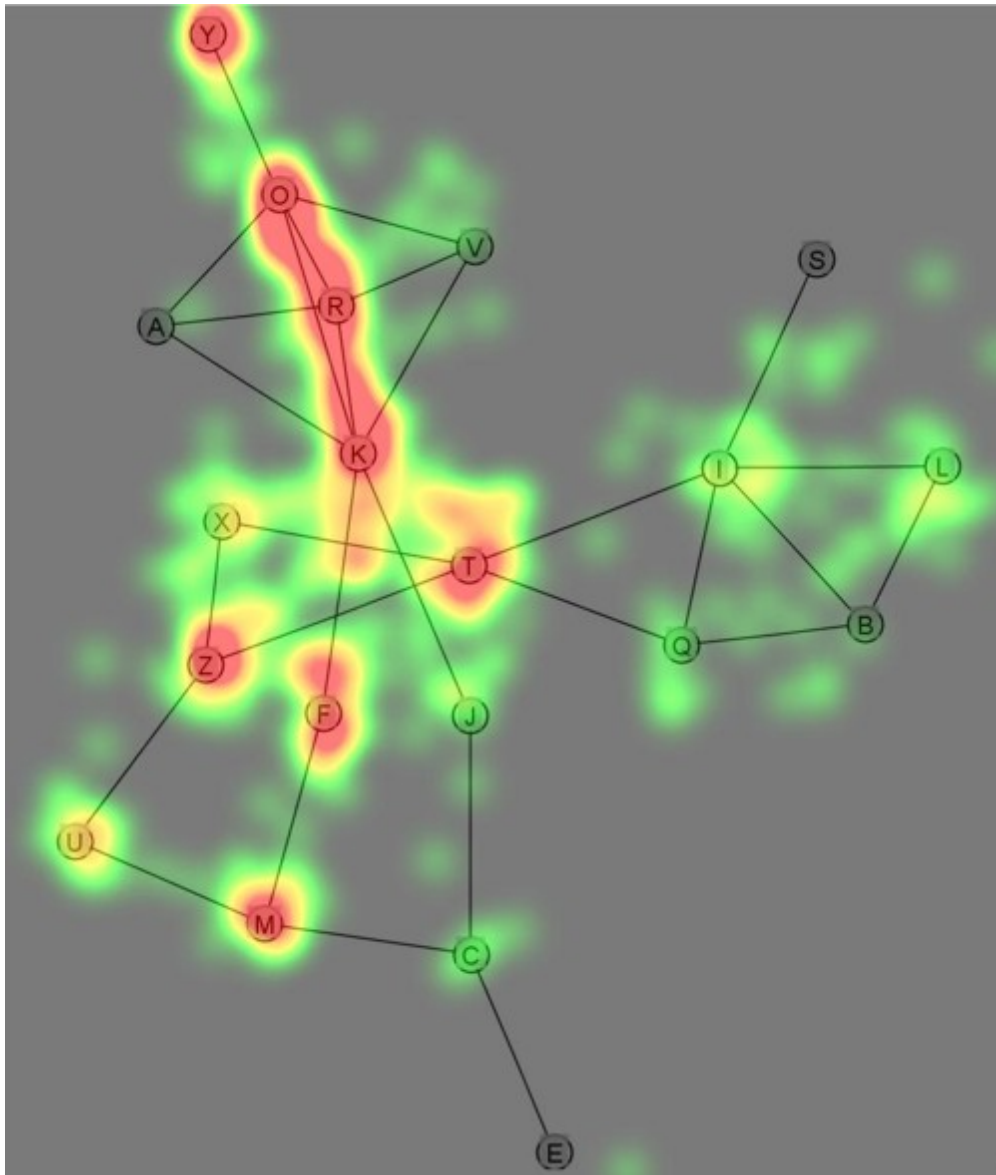


Abbildung 3.24: Heatmap Task 2, 20 Knoten, kräftebasiert

Heatmap aufgeführt werden kann, werden nur Abbildungen näher betrachtet, die besonders deutlich entsprechende Thesen untermauern oder widerlegen. Der Scheuklappeneffekt soll zunächst am kräftebasierten Layout nachgewiesen werden, dazu dient Abbildung 3.24.

Bei diesem Task sollte durch die Probanden ein Weg vom Knoten „L“ zum Knoten „Y“ gefunden und angegeben werden. Zunächst fällt auf, dass die Heatmap zu dieser Aufgabe keine gleichmäßige Färbung aufweist, sondern sehr stark differenziert zwischen Knoten, an denen der weitere Weg „eindeutig“ ist und Knoten, wo es mehrere Möglichkeiten gibt, den Weg weiterzuführen. Anfängen von Knoten „L“, haben quasi alle Probanden den direkten Weg über „I“ nach „T“ verfolgt. Der Scheuklappeneffekt zeigt sich bereits hier besonders deutlich, wenn man den Knoten „S“ betrachtet, der keinerlei Fixationen aufweist. Aber auch die Knoten „Q“ und „B“ werden kaum beachtet, da diese offenbar intuitiv als Umweg wahrgenommen wurden. Generell kann festgestellt werden, dass bei dieser Anfangsphase der Aufgabe der Weg für alle Probanden sehr klar zu sein scheint, da auch die Knoten „L“ und „I“, die bei fast allen genannten Lösungen den ersten Wegabschnitt markierten, nur sehr kurz fixiert wurden. Dies untermauert wiederum die These, dass das kräftebasierte Layout das Lösen eines Wegproblems sehr gut unterstützt. Die erste große Fixation findet sich bei Knoten „T“: Dies liegt darin begründet, dass sich hier für die Probanden zwei Möglichkeiten ergeben, wie der Weg fortgeführt werden soll, nämlich nach „X“ oder nach „Z“. Außerdem suchen die Probanden nach einer Möglichkeit, direkt von „T“ in Richtung „Y“ zu gelangen, was durch die „Eruption“ an „T“ in Richtung „K“/„Y“ deutlich wird. An dieser Stelle lässt sich auch klar die Bedeutung von Kantenkreuzungen erkennen. Betrachtet man die Kreuzung zwischen den Kanten (F,K) und (T,X), kann man eine deutliche Fixationshäufung feststellen. Hier ist also seitens der Probanden ein erhöhter Arbeitsaufwand nötig, um diese Überkreuzung korrekt zu bewerten, was letztendlich auch gelingt. Dadurch wird die These gefestigt, dass die Minimierung von Kantenkreuzungen generell ein wichtiges Ästhetikkriterium darstellt, da bei einer Kreuzung höhere Konzentration durch den Probanden gefordert ist, beim kräftebasierten Layout haben die Versuchsteilnehmer allerdings weniger Probleme diese Kreuzungen aufzulösen, was die Korrektheitsanalyse belegt. Der weitere Weg führt dann in der Regel zum Knoten „Z“. Auch hier gibt es eine Besonderheit, denn hier muss der Weg in eine Richtung („U“) fortgesetzt werden, die entgegengesetzt zur eigentlichen Zielrichtung verläuft. Dadurch ist ein längeres Verharren auf „Z“ mit einer Fixationseruption in Richtung „K“/„Y“ zu erklären.

Im Knoten „U“ tritt das gleiche Phänomen auf wie am Anfang des Weges, die Richtung ist eindeutig und somit muss Knoten „U“ kaum fixiert werden. Die Probanden

setzen dann den Weg über „M“ und „F“ fort. Hier ist wieder ein sehr gutes Beispiel für den Scheuklappeneffekt beim kräftebasierten Layout zu erkennen, denn der Weg wird ab „M“ sehr zielgerichtet eingeschlagen, was zur Folge hat, dass die Knoten „C“ und „J“, die einen Umweg darstellen würden, gar nicht in Betracht gezogen werden. Der Knoten „E“ wird sogar von keinem einzigen Probanden fixiert, da in diesem Fall eindeutig ist, dass dieser Knoten nichts zur Lösung beiträgt. Dieser Umstand scheint den Probanden unterbewusst klar zu sein, denn die Nichtbeachtung eines Knotens kann nicht bewusst erfolgen. Hier ergibt sich also ein weiterer wichtiger Hinweis auf die Existenz des Scheuklappeneffektes. Im weiteren Verlauf des Weges gelangen die Probanden dann zur interessanten Knotengruppe „K“, „R“, „O“. An der intensiven Färbung dieser Passage kann man erkennen, dass die Teilnehmer an dieser Stelle den intensivsten Arbeitsaufwand zu leisten hatten: Hier verlaufen einige Kanten sehr dicht und in einem spitzen Winkel zueinander. Hier lässt sich ein weiteres bekanntes Ästhetikkriterium belegen, nämlich eine gewünschte Maximierung des Winkels zwischen Kanten. Aber auch ein erneut sehr deutlicher Beleg für die Existenz eines Scheuklappeneffektes ist festzustellen, denn die alternativen Wege von „K“ aus über „A“ oder „V“ fallen genau diesem Effekt zum Opfer, sie werden quasi nicht beachtet und weisen keine Fixationen auf.

Abbildung 3.25 zeigt die entsprechende Heatmap für ein orthogonales Layout bei Task 2. Hier sollte der Weg vom Knoten „O“ zum Knoten „Z“ angegeben werden. Bereits die quantitative Analyse hat hier ergeben, dass sich das orthogonale Layout aufgrund der minimalen Anzahl der Kantenkreuzungen zwar sehr zuverlässig dazu eignet, einen Weg zwischen zwei vorgegebenen Knoten zu finden, allerdings ist dieser Weg in den seltensten Fällen ein kürzester Weg bei uniformen Pfadkosten. Dies wurde darauf zurückgeführt, dass das orthogonale Layout große Schwächen in der Darstellung adjazenter Beziehungen zwischen Knoten hat. Im Kontext der qualitativen Analyse (und insbesondere bei der Suche nach einem Scheuklappeneffekt) kann dieser Umstand anders formuliert werden: Werden lange Kanten, die adjazente Knoten verbinden, vom Probanden als solche wahrgenommen oder werden diese überhaupt nicht beachtet? Beim ersten Betrachten von Abbildung 3.25 fällt zunächst auf, dass die Punkte starker Fixationen (rot eingefärbt) fast ausschließlich auf Knoten entfallen, wohingegen die Kanten (auch solche, die Teil eines Weges sind) kaum fixiert werden. Diese Beobachtung ist noch stärker ausgeprägt als es beim kräftebasierten Layout der Fall ist und untermauert die Eignung des orthogonalen Layouts zur Bearbeitung des Wegeproblems, vorausgesetzt das Ziel ist nicht das Finden eines kürzesten Weges bei uniformen Pfadkosten. Denn wenn der Proband die Kante nicht fixieren muss, dann deutet dies darauf hin, dass der Verlauf der Kante sehr leicht nachvollziehbar ist. Im Vergleich zu den Kanten sind allerdings alle (beachteten) Knoten sehr stark gefärbt, also werden diese häufig fi-

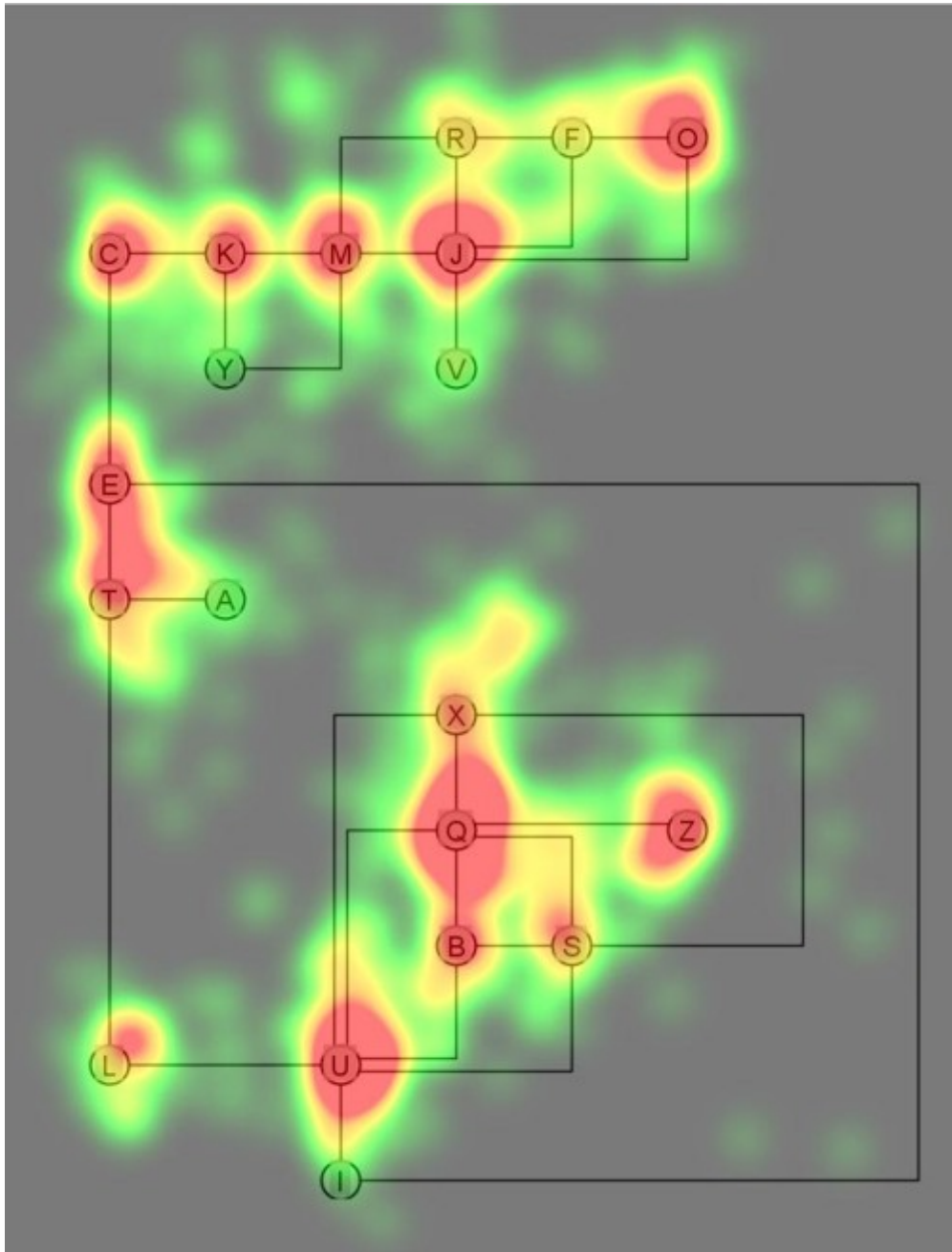


Abbildung 3.25: Heatmap Task 2, 20 Knoten, orthogonal

xiert. Hierbei ist auffällig, dass besonders Knoten mit einem hohen Ausgangsgrad ausgiebiger fixiert werden, so ist die Färbung der Knoten „U“, „Q“ oder auch „J“ jeweils deutlich stärker ausgeprägt als z.B. die Knoten „L“ oder „C“. Letztgenannte haben einen geringeren Ausgangsgrad, was gleichbedeutend mit der Tatsache ist, dass der weitere Weg für die Probanden schneller zu erfassen ist aufgrund von geringeren Auswahlmöglichkeiten. Hier wird allerdings ein Unterschied zum kräftebasierten Layout deutlich: während dort Knoten, bei denen klar ist, welche Kante als nächstes verfolgt wird, nur kurz fixiert werden, so sind die Fixationen beim orthogonalen Layout selbst bei eindeutigen Fällen (z.B. Knoten „L“) länger und auch häufiger.

Doch wie sieht es mit dem Scheuklappeneffekt aus? Genau wie beim kräftebasierten Layout können wir auch hier viele Belege dafür finden, dass bei dieser Aufgabenstellung tatsächlich ein solcher Effekt nachzuweisen ist. Startend bei Knoten „O“ nahmen viele Kandidaten den direkten Weg nach „J“, einige entschieden sich auch für den Umweg über „F“ (und „R“). Dann verlief bei allen Probanden der Weg geradlinig bis nach „C“. Auf diesem Weg können wir zwei Beispiele für „unwichtige“ Knoten finden, nämlich Knoten „V“ und Knoten „Y“, die nicht zur Lösung des Tasks benötigt werden. Folglich betrachtete auch kaum einer der Teilnehmer diese Knoten, sie wurden geradezu ausgeblendet. Ein ähnliches Verhalten legten die Probanden auch bei Knoten „A“ an den Tag, auch dieser konnte nichts zur Lösung beitragen und wurde folglich weitestgehend ausgeblendet. Während bei den gerade genannten Beispielen ein Ausblenden durchaus Sinn macht, da die Knoten tatsächlich in eine Sackgasse geführt hätten oder die Lösung zumindest ganz offensichtlich kompliziert hätten, so ist eine weitere Ausblendung aufgrund des Scheuklappeneffektes wesentlich gravierender und belegt einmal mehr die Unzulänglichkeiten des orthogonalen Layouts bei der Darstellung von Adjazenzen: Die Kante (E,I) hätte eine echte Verkürzung der Lösung dargestellt, allerdings wurde diese kaum beachtet. Dies spiegelt sich nicht nur in den genannten Lösungen wider, sondern auch bei Betrachtung von Abbildung 3.25. Denn wäre diese Kante von mehreren Probanden in Erwägung gezogen worden, so müssten zumindest an den Kantenknicken Spuren von Fixationen zu finden sein, aber auch das ist nicht der Fall. Man kann also davon ausgehen, dass der Scheuklappeneffekt nicht nur existiert, sondern auch ganz erheblichen Einfluss auf die Qualität der Lösung hat. Denn während er in den meisten Fällen de fakto unwichtige Knoten ausblendet (wie beim kräftebasierten Layout) und somit zur schnelleren Lösungsfindung beiträgt, so führt er auf der anderen Seite auch dazu, dass für die Lösung wichtige oder nützliche Kanten ausgeblendet werden und somit in kaum einer Lösung auftauchen.

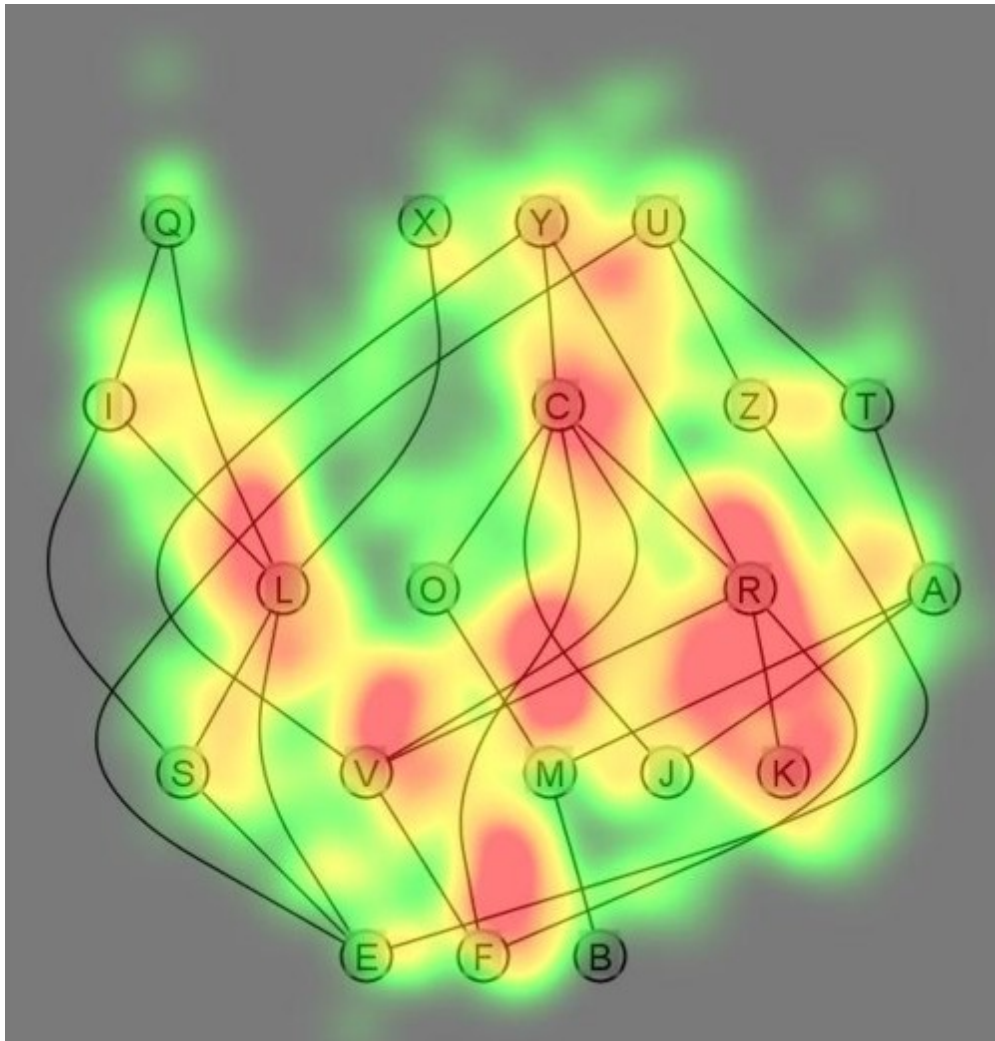


Abbildung 3.26: Heatmap Task 2, 20 Knoten, hierarchisch

Eine analoge Analyse für das hierarchische Layout basiert auf Abbildung 3.26. Bei dieser Aufgabe sollten die Probanden einen Weg von Knoten „Q“ zum Zielknoten „K“ finden. Bei dieser Abbildung zeigt sich jedoch bereits beim ersten Blick, dass die Färbung der Heatmap eine völlig andere Charakteristik aufweist als bei den anderen beiden Layouts. Damit wird die Sonderstellung, die das hierarchische Layout bereits in der quantitativen Analyse bei Task 2 eingenommen hat auch in der qualitativen Analyse mehr als deutlich. Im Gegensatz zu den beiden anderen Layouts, bei denen die Fixationen hauptsächlich bei Knoten zu finden sind, werden die Knoten hier kaum beachtet. Nur sehr wenige Knoten (z.B. „K“, „R“, „L“) weisen eine

intensive Rotfärbung auf, der Großteil der Knoten hingegen ist nur teilweise oder schwach gefärbt. Dennoch weist die Heatmap große Bereiche mit einer intensiven Rotfärbung auf, allerdings beim hierarchischen Layout liegt der Fokus eindeutig auf den Kanten. Bei genauerer Betrachtung kann sogar festgestellt werden, dass insbesondere Kantenkreuzungen, deren Zahl beim hierarchischen Layout besonders hoch ist, durch eine ebenso hohe Zahl von Fixationen gekennzeichnet sind. Ein besonders gutes Beispiel für diese Aussage liefert der Bereich über Knoten „M“, denn hier kreuzen sich vier Kanten auf sehr engem Raum und zumeist in einem spitzen Winkel, so dass bereits in der quantitativen Auswertung die Vermutung nahelag, dass es hier zu Problemen seitens der Probanden kam. Dies wird hier nun eindeutig belegt, denn eine intensive Rotfärbung ist begründet in einer hohen Zahl von Fixationen, was wiederum auf eine intensive Betrachtung aufgrund von mangelnder Zugänglichkeit hinweist.

Man kann in Abbildung 3.26 noch weitere Belege hierfür finden. Eine ähnliche Färbung wie die gerade betrachtete weisen die Kanten rund um den Knoten „L“ auf. Auch hier kommt es oberhalb von „L“ zur Kreuzung von vier Kanten, deren Schnittwinkel zwar recht groß ist, allerdings verlaufen die Kanten recht eng zueinander. Ein weiterer Beleg dafür, dass Kantenkreuzungen ein wichtiges Ästhetikkriterium sind und insbesondere beim hierarchischen Layout große Probleme verursachen, ist der Bereich zwischen „R“ und „K“. Dieser wurde aus zweierlei Gründen häufig fixiert: Zum einen findet man auch hier wie in den Vorbetrachtungen vier Kanten, die sich kreuzen und somit für den Betrachter schwierig zu verfolgen sind, allerdings sagt die Färbung dieses Bereichs auch etwas über das Suchverhalten der Probanden aus. Aufgrund des hohen Schwierigkeitsgrades von Task 2 bei diesem Layout scheiterten die Probanden bei der Wegsuche von „Q“ ausgehend, woraufhin die meisten als Alternativstrategie eine Rückwärtssuche vom Knoten „K“ ausgehend versuchten. Dies belegen auch die Aufzeichnungen der einzelnen Versuchspersonen. Doch wie sieht es bei diesem Layout mit dem Auftreten eines Scheuklappeneffektes aus? Trotz einer insgesamt flächigeren Färbung kann dennoch festgestellt werden, dass auch hier der genannte Effekt auftritt, wenn auch weniger eindeutig als bei den anderen Layouts. Den deutlichsten Hinweis hierauf liefert der Knoten „B“, der von keinem der Probanden fixiert wurde, da dieser sicherlich kein Teil der Lösung ist, da er eine „Sackgasse“ darstellt. Dies muss den Probanden auch ohne eine nähere Betrachtung, also mehr oder minder unterbewusst, klar sein.

Aber auch hier werden für die Lösung der Aufgabe wichtige Kanten ausgeblendet, was für das Gesamt-Ergebnis fatale Folgen hat. Das schlechte Abschneiden des hierarchischen Layouts bei Task 2 lässt sich nämlich besonders an der Kante (E,U) festmachen. Wie die Heatmap zeigt, wurde diese Kante kaum wahrgenommen

(außer an Stellen, an denen sie andere Kanten kreuzt), obwohl sie ein essentieller Teil des kürzesten Weges von „Q“ nach „K“ und auch der anderen Lösungen ist. Die Gründe für das Ausblenden dieser Kante dürften ähnlich gelagert sein, wie beim orthogonalen Layout: Die Kante verläuft über fünf Hierarchieebenen und ist damit sehr lang. Kanten dieser Länge werden allerdings, wie bereits gesehen, von einem menschlichen Betrachter aufgrund des Scheuklappeneffektes weitestgehend ignoriert. In diesem konkreten Fall kommt noch erschwerend hinzu, dass die Kante (E,U) in entgegengesetzter Richtung zum Zielknoten „K“ verläuft. Durch die Kombination der hier nachgewiesenen Zugänglichkeitsprobleme durch Kantenkreuzungen und den ebenfalls festgestellten Scheuklappeneffekt erhöht sich der Schwierigkeitsgrad dieses Tasks beim orthogonalen Layout überproportional stark, wodurch die schlechten Ergebnisse der quantitativen Analyse belegt und erklärt werden können. Abschließend kann man festhalten, dass der Scheuklappeneffekt bei allen Layouts auftritt, allerdings unterschiedlich stark ausgeprägt ist. Während beim kräftebasierten Layout zielsicher unwichtige Knoten ausgeblendet werden, führt der Effekt beim orthogonalen Layout aufgrund dessen Eigenschaften zum Ignorieren von wichtigen Kanten und Knoten. Weniger stark ist der Effekt beim hierarchischen Layout ausgeprägt, in Kombination von langen Kanten und vielen Kantenkreuzungen führt er aber dennoch zum schlechtesten Ergebnis aller drei Verfahren.

Eine mit dem Scheuklappeneffekt eng verwandte Thematik ist die Bedeutung von langen Kanten für die Lesbarkeit durch einen menschlichen Betrachter. Hier hat das kräftebasierte Verfahren einen großen Vorteil gegenüber den konkurrierenden Layouts, denn aufgrund des „spring-embedder“-Ansatzes werden lange Kanten in der Regel vermieden (eine lange Kante entspräche einer sehr stark gespannten Feder, so dass in einem Gleichgewichtszustand diese Feder vermutlich in einen entspannteren Zustand übergehen würde). Anders sieht es, wie bereits oben angeführt, bei den orthogonalen und hierarchischen Abbildungen aus. Beim hierarchischen Layout in Abbildung 3.26 wurde bereits auf die nicht vorhandene Fixation bei der Kante (E,U) hingewiesen. Aber auch andere, zum Teil wichtige, Kanten werden weitestgehend ignoriert, so z.B. auch die Kante (Y,V), deren Endknoten vier Schichten auseinander liegen sowie die Kante (E,Z), die ebenfalls über vier Schichten verläuft. Dies demonstriert ein Grundproblem des hierarchischen Layouts: Da keine adjazenten Knoten auf der gleichen Hierarchieebene liegen können, steigt bei hohem Durchschnittsgrad die Zahl der Schichten stark an. Dies wiederum bedeutet, dass zum Teil sehr lange Kanten entstehen, die über mehrere Schichten hinweg verlaufen. Durch diese Länge der Kanten steigt wiederum die Wahrscheinlichkeit von Kantenkreuzungen, die besonders bei diesem Verfahren, aufgrund häufiger spitzer Schnittwinkel, ein Zugänglichkeitsproblem darstellen. Daher muss es ein Ziel des hierarchischen Algorithmus sein, die Anzahl der Ebenen zu minimieren, um so kür-

zere Kanten zu fördern. Dies hat allerdings möglicherweise wiederum zur Folge, dass die Abbildung in vertikaler Richtung gestaucht wirkt. Die Kanten müssen dadurch stärker geschwungen werden, was zur Folge hat, dass die Schnittwinkel der Kanten noch spitzer werden. Man hat hier also einen „Trade-Off“ zwischen der Anzahl der Hierarchieebenen (und damit der Länge der Kanten) und dem Schnittwinkel. Hier sollte folglich, um die Zahl der Irritationen zu minimieren, beim Algorithmus verstärkt auf die Minimierung von Kantenkreuzungen Wert gelegt werden, um eine Optimierung des „Trade-Offs“ zu erreichen.

Beim orthogonalen Layout in Abbildung 3.25 stellen lange Kanten ebenfalls ein Problem dar, was die Analyse eines möglichen Scheuklappeneffektes verdeutlicht hat. Die Konsequenzen für den orthogonalen Algorithmus lassen sich auf die Formel „Bessere Darstellung von Adjazenzen“ reduzieren. Das beste Beispiel liefert die Kante (E,I). Diese Kante wurde kaum wahrgenommen, was nicht weiter verwundert, wenn man deren Verlauf betrachtet. Vermutlich wäre eine Kantenführung links, parallel zur Kante (T,L) wesentlich sinnvoller gewesen, ähnlich wie es bei den Kanten (U,Q) und (U,X) realisiert ist. Hier sollten also am Algorithmus Optimierungen vorgenommen werden, die eine Minimierung der Kantenlängen bei gleichzeitiger Kreuzungsminimierung zur Folge haben. Das Beispiel (E,I) zeigt recht deutlich, dass hier noch Optimierungsspielräume in der Einbettung der Kanten in das orthogonale Gitter vorhanden sind.

Eine weitere Vermutung aus der quantitativen Analyse, die es hier zu belegen gilt, lautet, dass bei den Tasks 3 und 4 die erste Strategie der Probanden darin bestand, zunächst eine „Form“ zu suchen, die dem vorgegebenen Teilgraphen entsprechen könnte und erst dann die eigentlich wichtigen adjazenten Beziehungen im Einzelnen abzuprüfen. Abbildung 3.27 liefert die hierzu passende Heatmap, die die Fixationen der Probanden bei der Bearbeitung von Task 3 zeigt. Eine andere Formulierung für die gesuchte „Form“ eines Teilgraphen oder einer Clique soll der Begriff „Kandidatengruppe“ von Knoten sein. Bei der betrachteten Abbildung lassen sich leicht alle von den Probanden gefundenen Kandidatengruppen erkennen: zum einen die Gruppe (F,R,J,M,K), dann (U,B,Q,X,S) und schließlich zogen einige Versuchsteilnehmer noch die Gruppe (E,T,A) als möglichen Teil einer Lösung in Erwägung. Die Kumulation der Fixationen in den verschiedenen Bereichen der Abbildung untermauert ganz eindeutig die These, dass zunächst nach Formen gesucht wird und erst in einem zweiten Schritt die eigentliche Überprüfung der Eigenschaften stattfindet. Hierdurch findet man auch eine Erklärung für das schlechte Ergebnis des orthogonalen Layouts bei den Tasks 3 und 4. Denn dadurch, dass bestimmte Knotengruppen sehr eng beieinander angeordnet werden, während andere adjazente Knoten durch sehr lange Kanten mit großem Abstand abgebildet werden, gibt die orthogonale

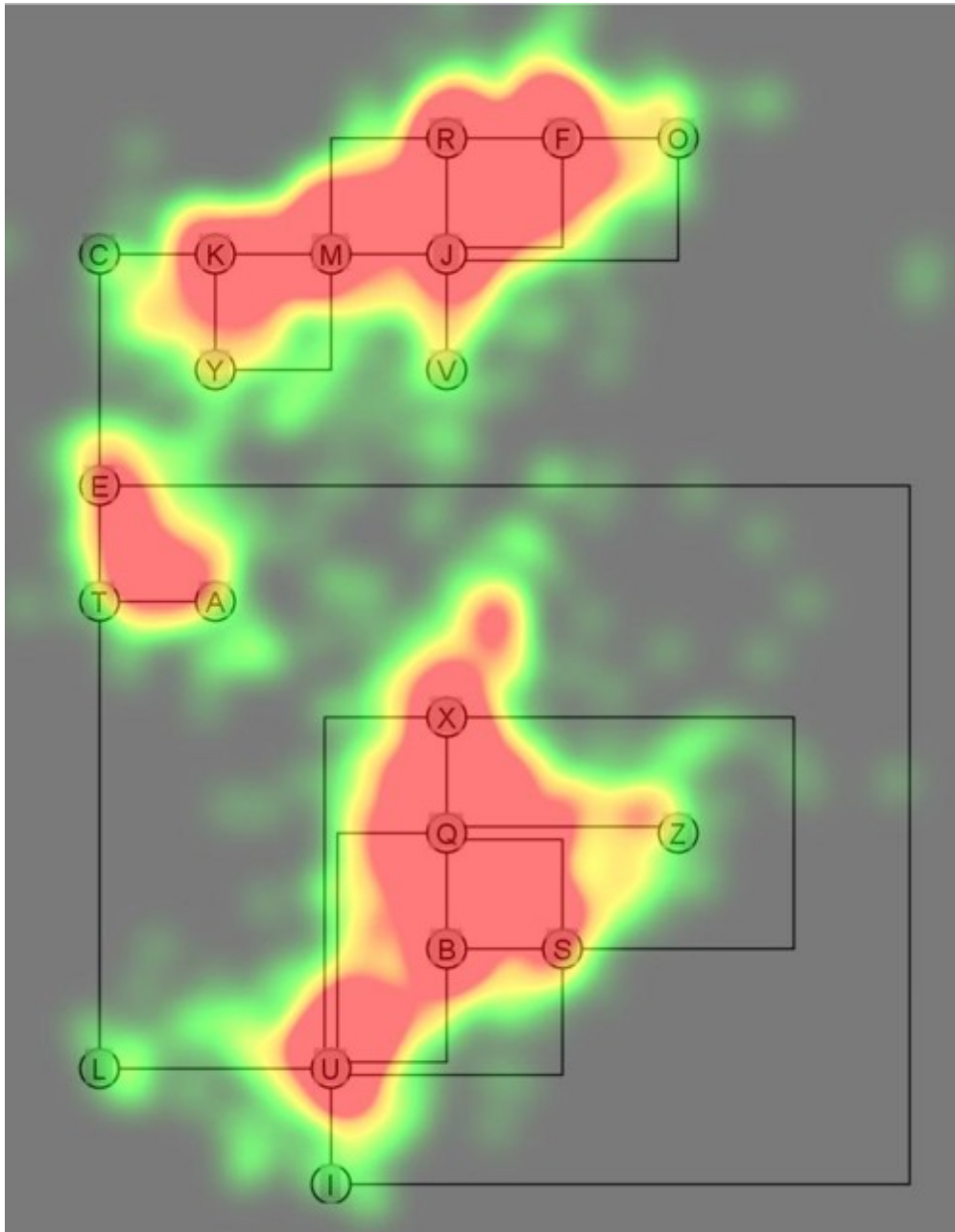


Abbildung 3.27: Heatmap Task 3, 20 Knoten, orthogonal

Abbildung bereits einige Kandidatengruppen vor und zerreit zugleich andere Knotengruppen durch zu lange Kanten.

Einen Beleg dafr liefert zum Beispiel die Kante (X,S). Vergleicht man die Lnge dieser Kante mit den Lngen der anderen Kanten dieser Knotengruppe, so wird diese Kante deutlich benachteiligt. Dies spiegelt sich auch in der Heatmap wider, denn diese Kante besitzt keine Fixationen, selbst an den Knoten gibt es keine „Eruptionen“ in Richtung der Kante. Dabei ist diese Kante ein wichtiger Teil einer mglichen Lsung (Q,B,U,S,X) zu Task 3. Die Kante (E,I), als weiteres Beispiel, wurde wie bei Task 2 auch hier fast vllig ignoriert, der Knoten „I“ wurde dabei generell kaum beachtet. Diese Ergebnisse belegen die Probleme des orthogonalen Layouts fr die Darstellung von stark Adjazenz-bezogenen Zusammenhngen. Denn das worst-case - Verhalten dieses Verfahrens wird jede mgliche Lsung durch lange Kanten so weit auseinanderziehen, dass eine logische Verknpfung durch einen menschlichen Betrachter kaum noch mglich ist.

Ein hnlich schlechtes Ergebnis wie das orthogonale Layout erreichte in der quantitativen Analyse das hierarchische Layout bei den Tasks 3 und 4. Die passende Heatmap dazu liefert Abbildung 3.28. Auch hier wurde die Vermutung aufgestellt, dass die adjazenten Beziehungen von Knoten durch die Aufteilung auf die verschiedenen Hierarchieebenen zerstrt werden und somit die Hauptstrategie der Probanden, nmlich nach einer bestimmten Form zu suchen, nicht zum Erfolg fhren konnte. Denn besonders die Darstellung einer Clique wird durch das hierarchische Layout stark verzerrt, so dass alle beteiligten Knoten auf einer eigenen Hierarchieebene abgebildet werden. Diese Praxis hat also nicht nur zur Folge, dass lange Kanten mit vielen Kreuzungen entstehen, sondern die Form wird auch vertikal verzerrt. Ein gutes Beispiel stellt die mgliche Lsung (X,Q,V,Y) dar. Die uersten Knoten der Clique haben einen Abstand von fnf Hierarchieebenen, wodurch die Kante (X,Y) ber die gesamte Hhe der Abbildung verlaufen muss. Das Resultat ist eindeutig: Aufgrund dieser unzugnglich langen Kanten (neben (X,Y) hat der Knoten noch eine zweite lange Kante (X,V) als Anbindung) wird der Knoten „X“, obwohl essentieller Bestandteil der Lsung, kaum beachtet. Die Kandidatengruppen sind anders aufgeteilt. Obwohl „R“ nur minimal ber eine einzige Kante mit der Gruppe verbunden ist, so wird er aufgrund seiner rumlichen Nhe dennoch eher als Lsungs - Kandidat betrachtet, als der gut angebundene, aber rumlich nicht zu dieser Knotengruppe gehrende, Knoten „X“. Generell kann man feststellen, dass die stark fixierten Bereiche, die Kandidatengruppen, beim hierarchischen Layout noch strker ausgeprgt, das heit noch raumgreifender, sind. Die Hufung der Fixationen erklrt sich daraus, dass nach der Auswahl einer Kandidatengruppe die Beziehungen der Knoten zueinander berprft werden mssen. Doch whrend die Kanten

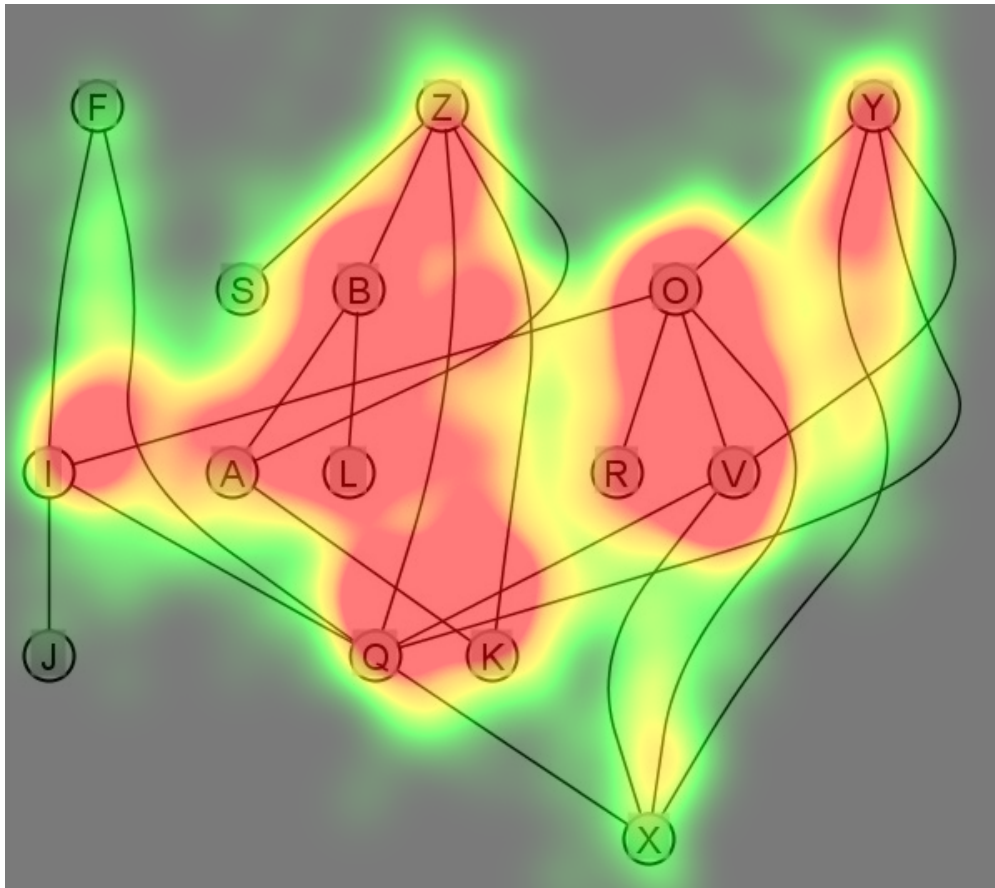


Abbildung 3.28: Heatmap Task 4, 15 Knoten, hierarchisch

beim orthogonalen Layout zumindest leicht verfolgt werden können, so kommt beim hierarchischen Layout erschwerend hinzu, dass dort zusätzlich große Schwierigkeiten auftreten, die entsprechenden Kanten zu verfolgen. Dies wiederum bedeutet eine größere Anstrengung und führt so zwangsläufig zu häufigeren und längeren Fixationen.

Nachdem die Probleme des orthogonalen und des hierarchischen Layouts nun belegt sind, ist es auch möglich nachzuweisen, dass die vermutete Stärke des kräftebasierten Layouts, nämlich das Erhalten einer bestimmten äußeren Form, beziehungsweise bei Cliques das Zuweisen von sehr charakteristischen Knotenzusammenstellungen, tatsächlich der Hauptgrund für das gute Ergebnis dieses Verfahrens ist? Zur Analyse dieser Thematik dient Abbildung 3.29, die beispielhaft die Fixationen aller Probanden für Task 4 bei einem kräftebasierten Layout demonstriert. Hier wird die

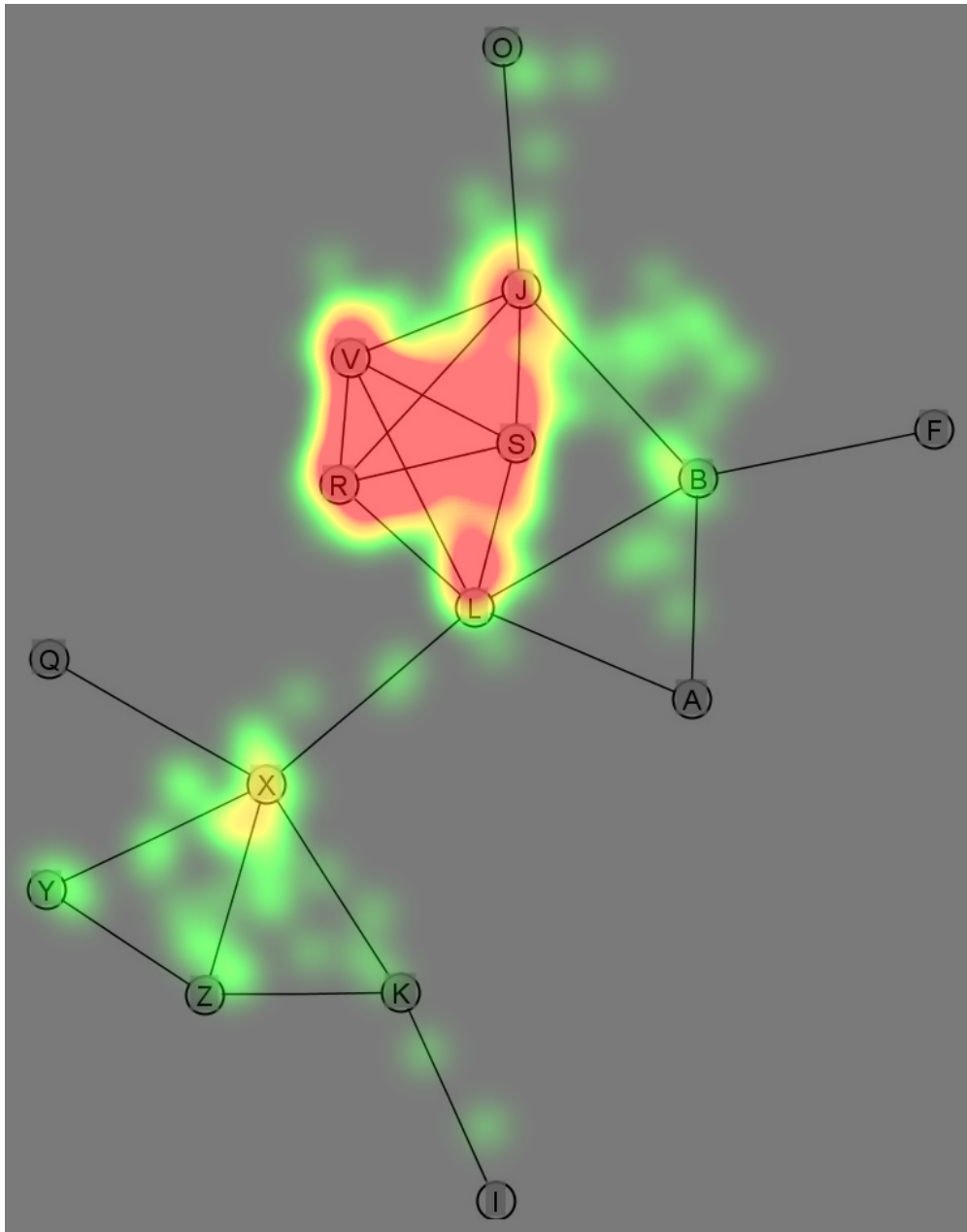


Abbildung 3.29: Heatmap Task 4, 15 Knoten, kräftebasiert

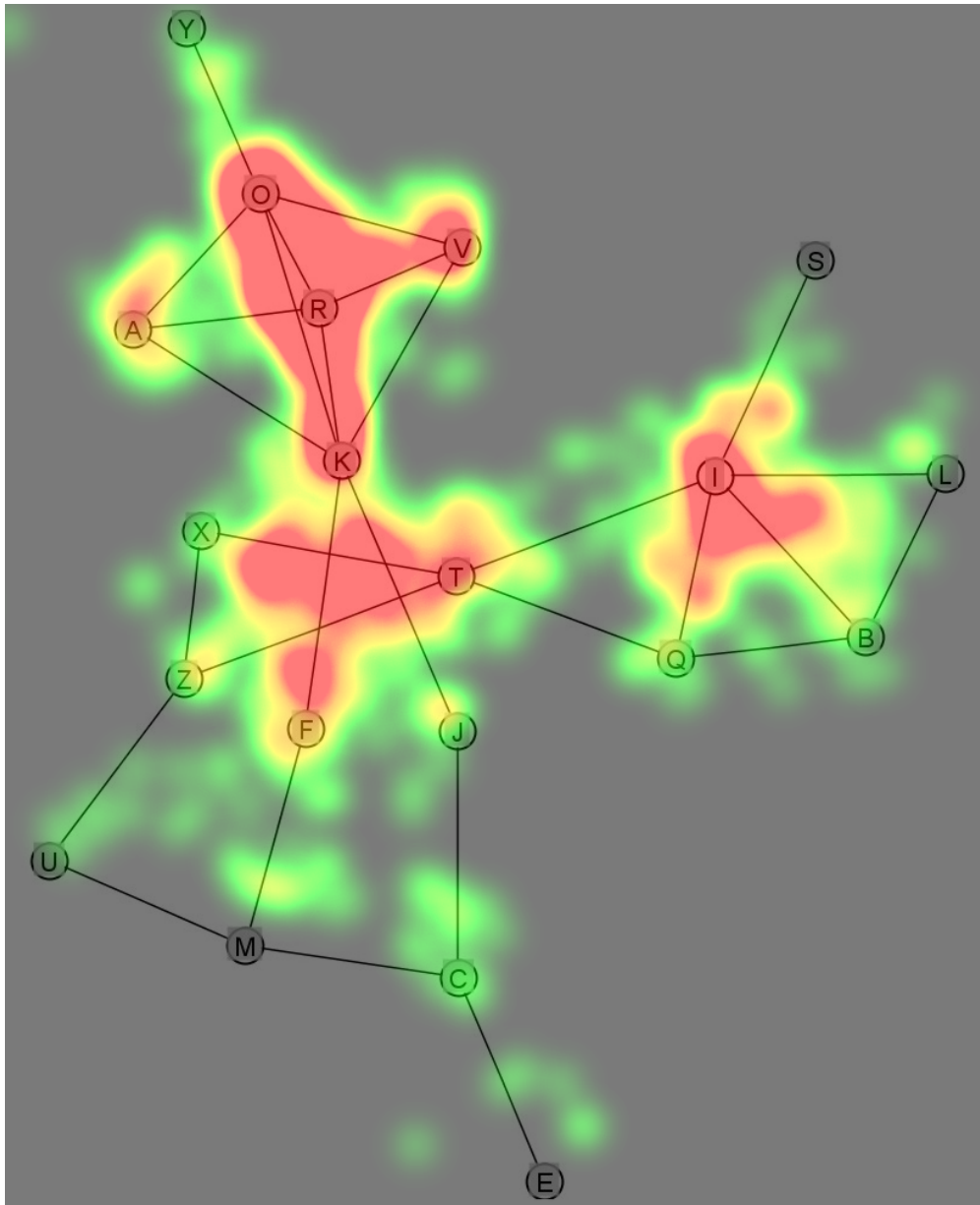


Abbildung 3.30: Heatmap Task 4, 20 Knoten, kräftebasiert

Strategie der Probanden sehr deutlich. Es finden kaum Fixationen außerhalb der Kandidatengruppe (J,V,R,S,(L)) statt, was beweist, dass diese Gruppe von Knoten für die Versuchsteilnehmer von vorneherein ohne großen Suchaufwand als die wahrscheinlichste Lösung angesehen wurde. Eine geringe Zahl von Fixationen weist auch noch die Knotengruppe um „X“ auf, allerdings war hier den Probanden sehr schnell klar, dass es sich bei dieser Gruppe nicht um eine Clique handelt, was wiederum für die hervorragende Zugänglichkeit des Layoutverfahrens spricht. Auch der Umstand, dass alle weiteren Knoten zur Lösung des Problems überhaupt nicht in Betracht gezogen wurden, untermauert diese Erkenntnis.

Nun ist allerdings die Clique in Abbildung 3.29 sehr eindeutig und schnell zu erkennen. Wie sieht es aus, wenn das kräftebasierte Verfahren die Clique weniger „charakteristisch“ abbildet? Dies ist in Abbildung 3.30 dokumentiert. Die Lösung zu Task 4, nämlich die Knotengruppe (O,A,R,K) ist weniger deutlich auf den ersten Blick zu erkennen, da das Schema „Raute mit einem Kantenkreuz in der Mitte“ hier etwas verzerrt wurde. Trotzdem zeigt die Heatmap, dass die Probanden ihre Strategie nicht ändern, sondern trotzdem nach Knotengruppen suchen, die eine ähnliche „Form“ haben, wie sie der Idealvorstellung einer 4er-Clique entspricht. So wird z.B. die Gruppe (I,L,B,Q) als möglicher Kandidat in Erwägung gezogen (vergleiche besonders die „Eruptionen“ um den Knoten „I“), da diese eine Rautenform aufweist, allerdings wird diese Gruppe aufgrund des fehlenden Kantenkreuzes schnell wieder verworfen. Kantenkreuzungen weist wiederum die Gruppe (T,X,Z,F,J) auf, was zur Folge hat, dass die entsprechenden Knoten und besonders die beteiligten Kanten intensiv fixiert werden, allerdings wurde auch diese Gruppe von den meisten Versuchsteilnehmern schnell verworfen. Bleibt als letzte Kandidatengruppe nur noch die korrekte Lösung, wobei hier auch noch der Knoten „V“ in Erwägung gezogen wird, er hat sogar einen offenbar stärkeren Bezug zur Gruppe (O,R,K) als der Knoten „A“, was in diesem Fall an der räumlichen Nähe liegt, denn wie bereits erwähnt bildet das kräftebasierte Verfahren adjazente Knoten nahe beieinander ab. Dennoch lieferte dieser Graph, trotz seiner nicht ganz offensichtlichen Lösung, ein sehr gutes Ergebnis und demonstriert die Stärke des kräftebasierten Layouts bei stark Adjazenz-bezogenen Aufgaben.

Eine weitere These, die im Vorfeld der Evaluation betrachtet wurde, ist ein möglicher Vorteil des hierarchischen Layouts bei Task 1, der auf folgenden Annahmen beruht: Zum einen ist die hierarchische Darstellung eines Graphen in der Regel kompakter als bei den anderen Verfahren, das heißt der Proband hat mehr Knoten in seinem Blickfeld und der Blick muss weniger wandern. Zusätzlich ordnet das Layout die Knoten in übersichtlichen Hierarchieebenen an, so dass eine systematische Suche von oben nach unten möglich ist. Doch wählten die Versuchsteilnehmer

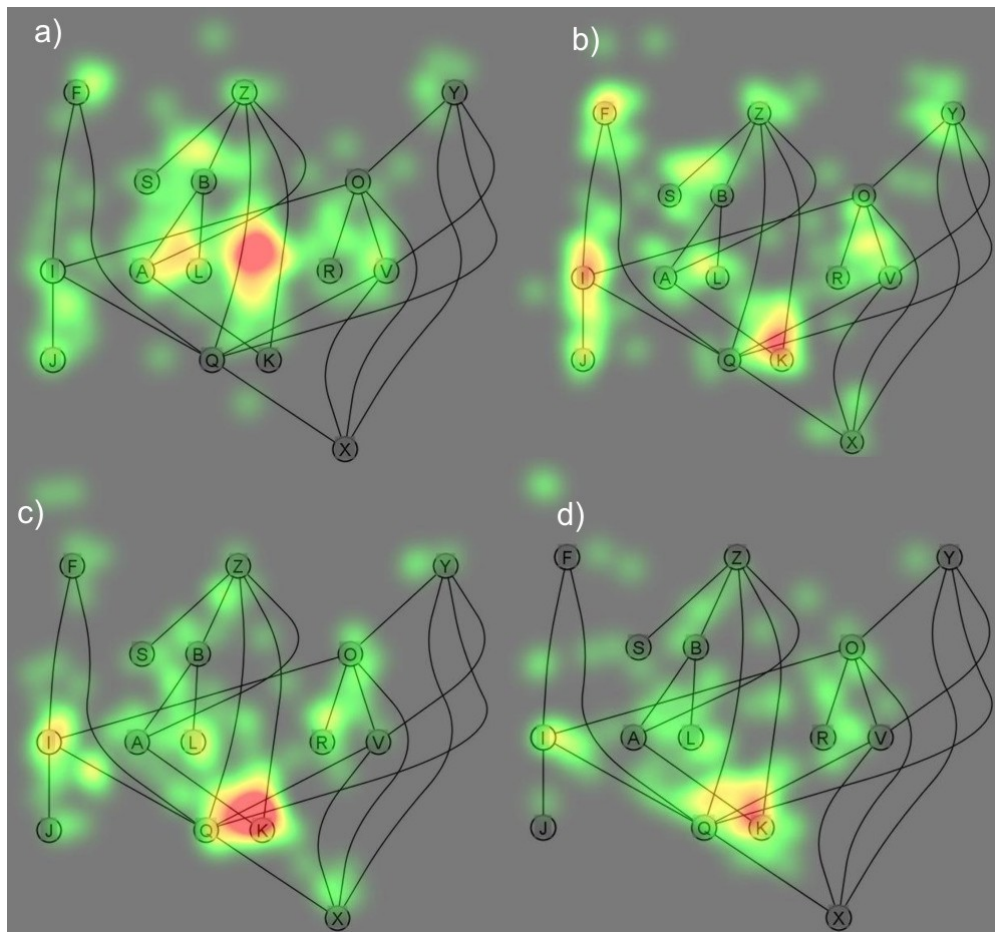


Abbildung 3.31: Heatmap Task 1, 15 Knoten, hierarchisch (4 Phasen)

tatsächlich diese Strategie, um bei Task 1 zu einem möglichst schnellen Ergebnis zu gelangen, oder wurde hier ein anderes Suchverfahren an den Tag gelegt? Die quantitative Analyse liefert zu dieser Frage bereits eine eindeutige Antwort, die lautet, dass keines der verwendeten Verfahren einen messbaren Vorteil bei der Suche nach einem speziellen Knoten liefert. Diese Tatsache wurde damit begründet, dass hier jeder Knoten isoliert betrachtet wird, also adjazente Beziehungen, deren Darstellung den Hauptunterschied der Verfahren ausmachen, keine Rolle spielen. Ob die Suchstrategien tatsächlich gleich sind, kann man erkennen, wenn man die Abbildungen 3.31 und 3.32 betrachtet. Bei diesen Abbildungen werden jeweils die ersten vier Sekunden der Aufzeichnungen aufgesplittet und von a) bis d) getrennt aufgeführt, das heißt jede Teilabbildung ist eine Sekunde der Aufzeichnung. Aufgrund der sehr kurzen Zeit sind die zu erkennenden Fixationsfärbungen natürlich wesentlich

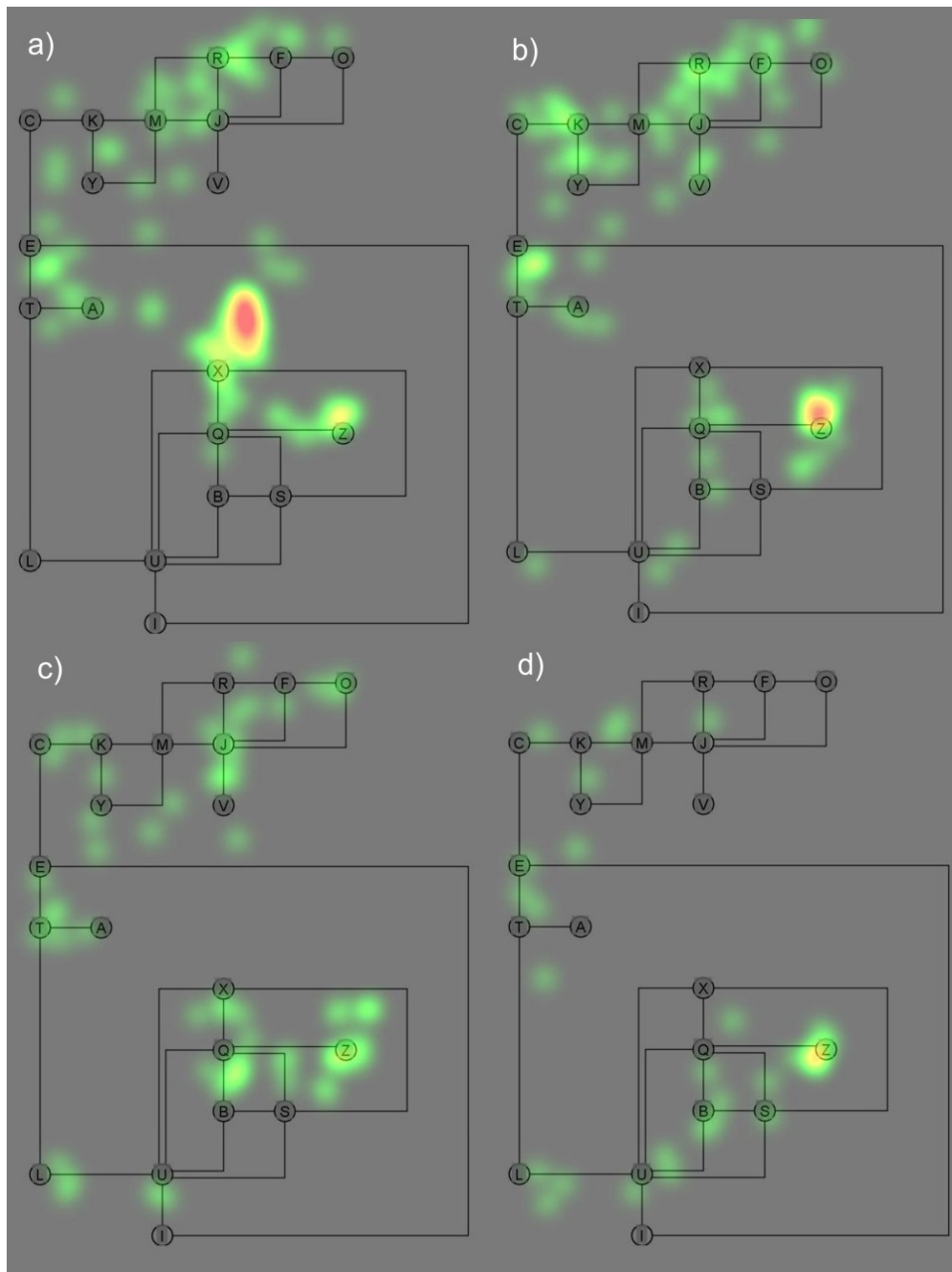


Abbildung 3.32: Heatmap Task 1, 20 Knoten, orthogonal (4 Phasen)

schwächer ausgebildet als bei den bisherigen Grafiken. Dies war allerdings aufgrund der sehr kurzen Durchschnittsbearbeitungszeiten bei Task 1 nötig.

Bei Abbildung 3.31 a) fällt zunächst die starke Färbung im Zentrum der Grafik auf, diese entstand allerdings durch das vorher eingeblendete Fixationskreuz, wodurch der Blick innerhalb der ersten Sekunde nach Umschalten auf den Graphen noch auf dieser Stelle fixiert bleibt. Die weiteren Fixationen haben jedoch alle gemeinsam, dass sie sich im linken oberen Abschnitt von dieser Hauptfixation befinden. Die Knoten im rechten unteren Bildbereich werden ignoriert („K“ (gesuchter Knoten), „X“, „Y“). Diese Beobachtung verstärkt sich noch mit Teilabbildung b), denn jetzt wird deutlich, dass die Suche der Probanden tatsächlich auf der linken oberen Seite beginnt, was man an den Färbungen der Knoten „F“, „I“, „J“ oder auch von „Z“ und dem gerade noch nicht betrachteten Knoten „Y“ erkennen kann. Erste Probanden haben auch „K“ bereits gefunden. Diese Verschiebung der Fixationen in „süd-westlicher“ Richtung setzt sich auch in c) weiter fort, bis schließlich in d) der Knoten „K“ von fast allen Teilnehmern gefunden ist. Die Frage, wie die Probanden einen bestimmten Knoten suchen, kann also beantwortet werden: Die Suche verläuft von links oben nach rechts unten, also genau wie man eine Buchseite lesen würde. Die Unabhängigkeit vom Layoutverfahren zeigt sich in Abbildung 3.32, denn das Verhalten der Probanden ist absolut identisch wie gerade betrachtet: In a) finden wir zunächst wieder die „Nachfixationen“ des Fixationskreuzes in der Mitte des Graphen. Allerdings werden auch hier die ersten Fixationen im linken oberen Teil ersichtlich, während im rechten unteren Teil des Bildes keinerlei Fixationen zu finden sind. In b) wird diese gedachte „Fixationslinie“ nochmals deutlicher erkennbar und bewegt sich in „süd-östlicher“ Richtung vorwärts. Ein Teil der Probanden hat den gesuchten Knoten „Z“ auch bereits gefunden. Über c) nach d) wird das weitere Wandern der Fixationen nach rechts und unten fortgesetzt, bis schließlich alle Teilnehmer „Z“ gefunden haben. Auch hier lässt sich das gleiche Suchmuster erkennen wie beim hierarchischen Layout, wodurch die These widerlegt ist, dass dieses aufgrund seiner hierarchischen Struktur einen Vorteil bei der Suche bestimmter Knoten hätte.

Die Betrachtung von Heatmaps zu Task 5 soll nun belegen, dass auch diese Aufgabenstellung, wie bereits in der quantitativen Analyse vermutet, als layoutunabhängig betrachtet werden kann. Die Abbildungen 3.33 bis 3.35 liefern hierzu die Grundlage. Bei Task 5 gilt es zu beachten, dass diese Aufgabe von ihrer Grundausrichtung her \forall -quantifiziert ist. Das bedeutet, dass ein Proband alle Knoten betrachten muss, um sich sicher zu sein, die korrekte Lösung auf die Frage nach dem höchsten Grad zu finden. Doch bereits die quantitative Analyse hat gezeigt, dass die Bearbeitungszeiten bei diesem Task sowohl unanständig vom verwend-

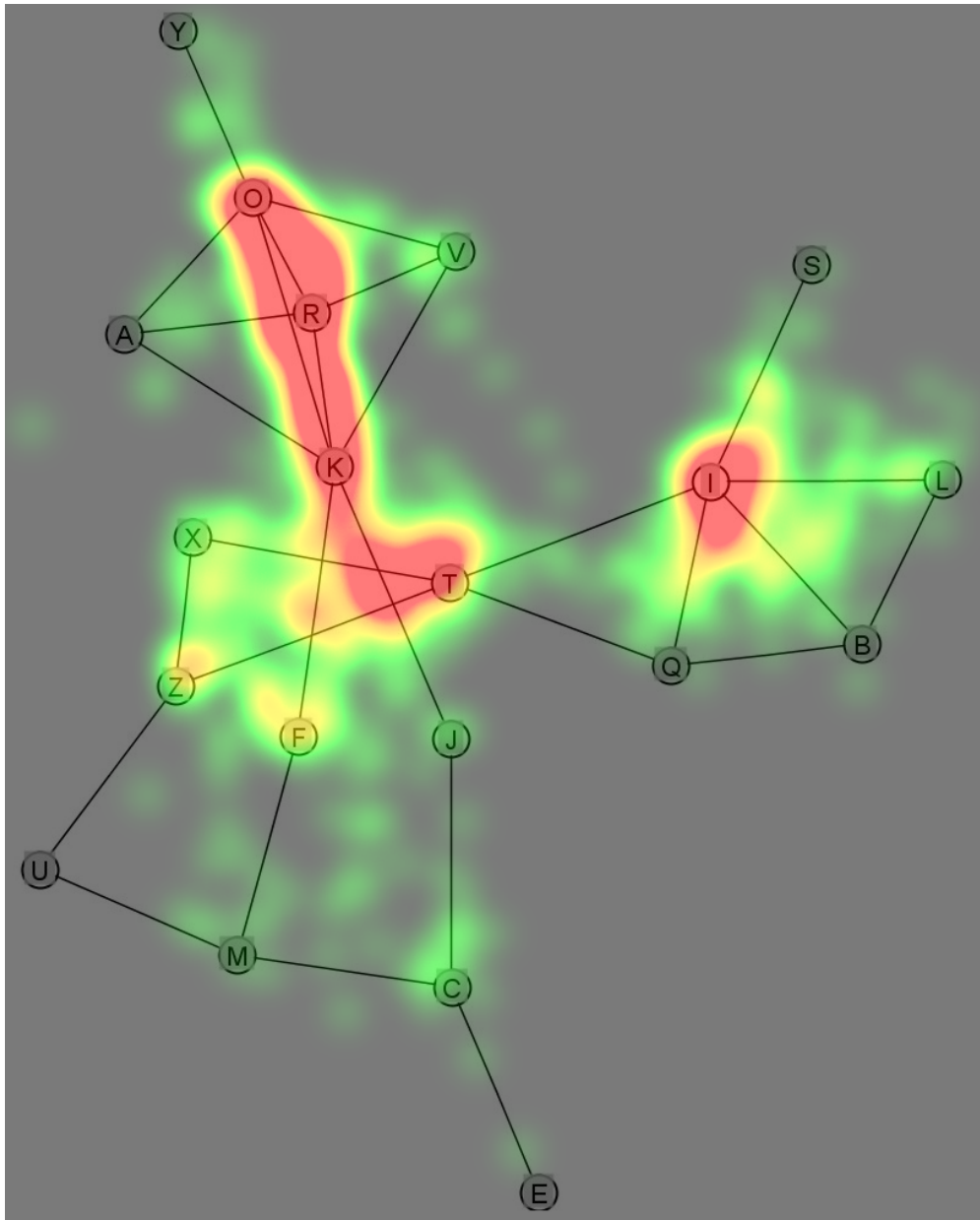


Abbildung 3.33: Heatmap Task 5, 20 Knoten, kräftebasiert

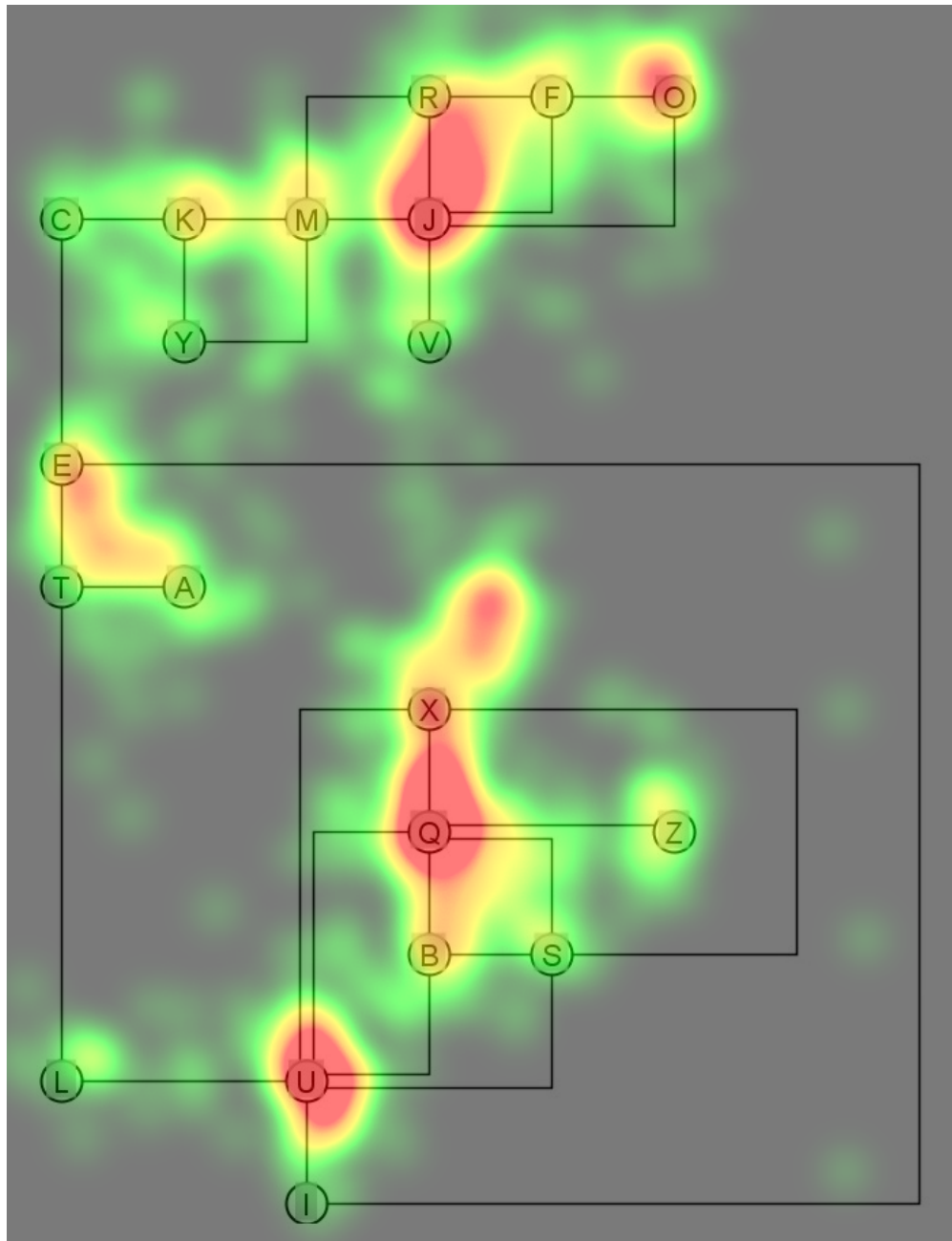


Abbildung 3.34: Heatmap Task 5, 20 Knoten, orthogonal

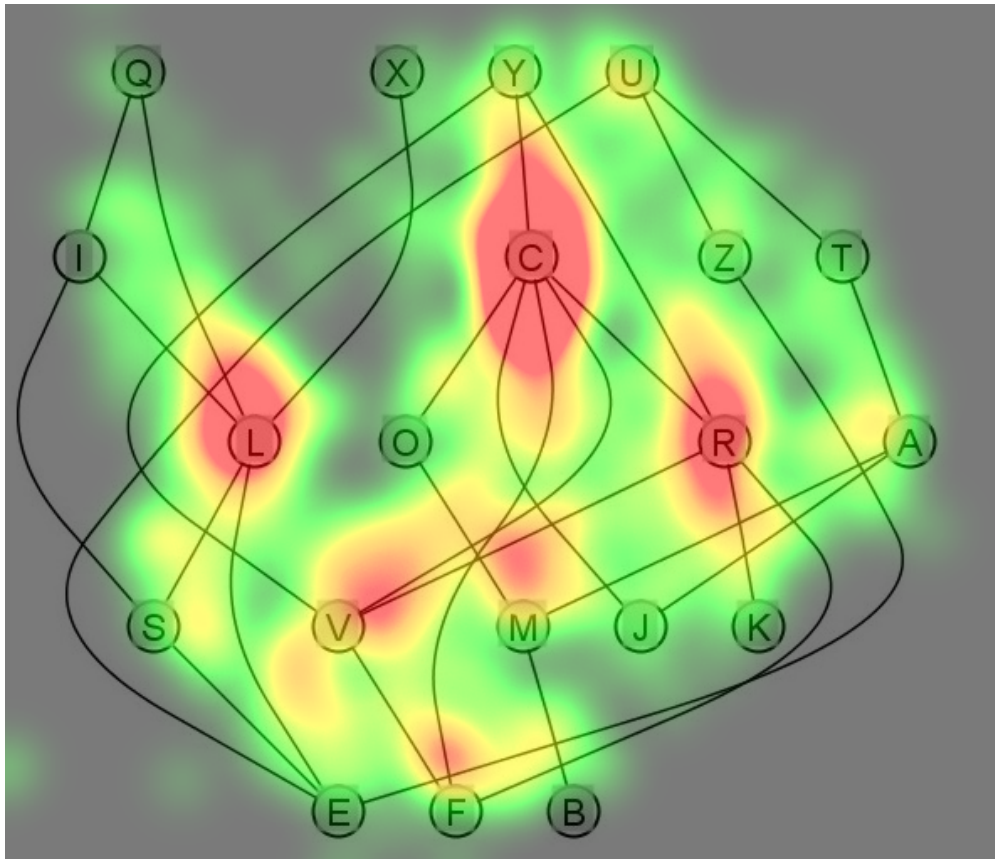


Abbildung 3.35: Heatmap Task 5, 20 Knoten, hierarchisch

ten Graphlayout als auch weitestgehend von der Komplexität des Graphen sind. Besonders aus letzterem muss man schließen, dass die Versuchsteilnehmer keineswegs alle Knoten betrachten, sondern ihre Suche auf einige wenige Kandidaten - Knoten beschränken und dabei auch nur die unmittelbare Umgebung dieser Knoten betrachten. Damit spielen adjazente Beziehungen, und somit das verwendete Graphlayout, eine untergeordnete Rolle. Abbildung 3.33 zeigt eine Heatmap eines kräftebasierten Graphen. Hierbei bestätigen sich die aufgestellten Thesen bereits sehr deutlich. Die Abbildung ist nur sehr partiell rot gefärbt, mit anderen Worten wurden nur sehr wenige Knoten tatsächlich fixiert. Man kann sogar soweit gehen zu sagen, dass ein Großteil der Knoten vollkommen ignoriert wurde. Diese Knoten kamen also bei den Betrachtern überhaupt nicht als Kandidat für den höchsten Grad in Betracht. Tatsächlich intensiv betrachtet wurden hier nur fünf von zwanzig Knoten, nämlich „O“, „R“, „K“, „T“ und „I“. Warum diese Knoten? Warum die anderen nicht? Die Erklärung liegt auf der Hand: Alle diese Knoten haben gemein-

sam, dass ihre unmittelbare Umgebung stark von Kanten frequentiert wird, sei es durch ausgehende Kanten oder (wie bei „R“ und „T“) durch kreuzende Kanten in unmittelbarer Knotennähe. Dadurch erscheint der Bereich um die Knoten herum als „dunkler“ als bei den anderen möglichen Kandidaten. Die Probanden konzentrieren sich also nur auf diese wenigen Knoten. Es wird weiter deutlich, dass die Teilnehmer auch nur die wirklich unmittelbare Umgebung um einen Kandidaten betrachten, die zum Kandidaten adjazenten Knoten werden hingegen nicht geprüft. Dies kann sehr gut am Knoten „I“ nachvollzogen werden, dessen Nachbarknoten „Q“, „B“ und „L“ keinerlei Fixationen aufweisen. Für das kräftebasierte Verfahren treffen die Thesen also in vollem Maße zu, um tatsächlich von einer Layoutunabhängigkeit sprechen zu können, müssen auch die beiden anderen Layouts entsprechende Hinweise liefern.

Abbildung 3.34 liefert daher die Heatmap für das entsprechende orthogonale Layout. Die Beobachtungen, die soeben beim kräftebasierten Layout gemacht wurden, lassen sich hier fast identisch ebenfalls nachvollziehen. Auch hier gibt es Kandidatenknoten, die intensiv fixiert wurden („Q“, „U“, „J“), während die meisten anderen Knoten weitaus weniger Beachtung fanden. Ein Unterschied zum kräftebasierten Layout ist allerdings, dass fast jeder Knoten zumindest rudimentär fixiert wurde. Ein ähnliches Verhalten der Probanden war auch schon bei Task 2 zu erkennen, denn während dort beim kräftebasierten Ansatz einige Knoten, die Teil der Lösung waren, kaum fixiert wurden, so wurde beim orthogonalen Layout jeder Knoten auf dem Weg intensiv fixiert. Diese Intensität erreichen hier, wie bereits festgestellt, nur die Kandidatenknoten. Genau wie in Abbildung 3.33 wird auch hier nur deren unmittelbare Umgebung betrachtet und nicht die benachbarten Knoten analysiert, wie man zum Beispiel bei Knoten „U“ deutlich feststellen kann.

Das hierarchische Layout in Abbildung 3.35 liefert nun die abschließende Bestätigung für die Theorie der Layoutunabhängigkeit von Task 5. Trotz eines völlig unterschiedlichen Layoutansatzes im Vergleich zum kräftebasierten Verfahren sind die Parallelen nicht zu übersehen: Auch hier kristallisieren sich nur wenige echte Kandidaten heraus, nämlich die Knoten „L“, „C“ und „R“. Wie oben bereits bemerkt, suchten die Probanden gerade diese Knoten heraus, weil die unmittelbare Umgebung dieser Knoten stark von ausgehenden („R“ und „C“) und auch von kreuzenden („L“) Kanten gekennzeichnet ist. Ebenfalls von Kanten stark frequentierte Bereiche (oberhalb von „V“ und „M“) werden von einigen Teilnehmern ebenfalls kurz in Betracht gezogen, aufgrund der Erkenntnis, dass dieser „dunkle Bereich“ allerdings nicht in einer Beziehung zu den entsprechenden Knoten gehört (keine Fixationen von „M“ oder „V“ direkt), werden diese Knoten jedoch wieder schnell verworfen. Eine weitere Parallele zu den anderen Layoutsverfahren: Auch hier wer-

den die adjazenten Knoten nicht betrachtet, sondern ausschließlich die unmittelbare Umgebung der Kandidaten.

Zusätzliche Kurzstudie zu Task 5

Aufgrund der Beobachtungen bei der Heatmapanalyse von Task 5 wurde im Anschluss an die Auswertung der Hauptevaluation noch eine zusätzliche Studie realisiert, mit deren Hilfe geklärt werden sollte, wie schnell ein menschlicher Betrachter eines Graphen einen Kandidaten für den maximalen Knotengrad finden kann. Durch die Abbildungen 3.33 bis 3.35 wird bereits deutlich, dass die Wahrnehmung bei der Kandidatenfindung auf einige wenige Knoten beschränkt ist. Da weitestgehend keine Fixationen bei anderen Knoten zu finden sind, stellt sich die Frage, ob das Finden von Kandidaten und somit das Ausblenden der anderen Knoten unterbewusst geschieht, oder ein bewusster Prozess ist. Von einer unterbewussten Wahrnehmung kann man dann sprechen, wenn die Tasks in einem Zeitraum von ca. 200 bis 250 ms bearbeitet werden können. Man spricht hier von einer präattentiven Wahrnehmung [Hea06]. Zur Feststellung, ob tatsächlich eine präattentive Wahrnehmung des Knotens mit dem maximalen Grad vorliegt, wäre ein komplexer Versuchsaufbau von Nöten gewesen, bei dem unter anderem die Netzhaut der Probanden abgetastet werden müsste. Da hierzu weder die technischen Voraussetzungen noch die benötigte Zeit zur Verfügung standen, wurde ein alternativer Versuchsaufbau entworfen, durch den zwar keine eindeutige Unterscheidung zwischen präattentiver und bewusster Wahrnehmung möglich ist, allerdings können in den Ergebnissen Tendenzen gefunden werden, die im Folgenden präsentiert werden sollen.

Der Versuchsaufbau basiert auf einer eigens für diese Studie entworfenen HTML-Seite, bei der mit Hilfe von JAVASCRIPT die Abbildung eines Graphen für eine vorgegebene Zeitspanne eingeblendet werden kann und sich nach Ablauf dieser Spanne automatisch ausblendet. Hierzu wurden ebenfalls speziell für dieses Experiment sechs verschiedene Graphen per Hand entworfen, die als solche Einblendungen dienen sollen. Diese ungerichteten Graphen haben spezielle Eigenschaften: Alle bestehen aus 25 Knoten, jeder dieser Knoten hat den Grad „3“. Bei drei Graphen ist allerdings ein Knoten zu finden, der einen signifikant höheren Grad aufweist, dieser hat nämlich Grad „7“. Dadurch ist es bei ausreichender Betrachtungszeit sehr leicht möglich festzustellen, ob ein Knoten einen (deutlich) höheren Grad hat oder ob alle Knoten den gleichen Grad aufweisen. Ein Beispiel für einen solchen Graphen (mit einem höhergradigen Knoten) liefert Abbildung 3.36. Die HTML-Seite war nun derart implementiert, dass bei Tastendruck durch einen Probanden jeweils ei-

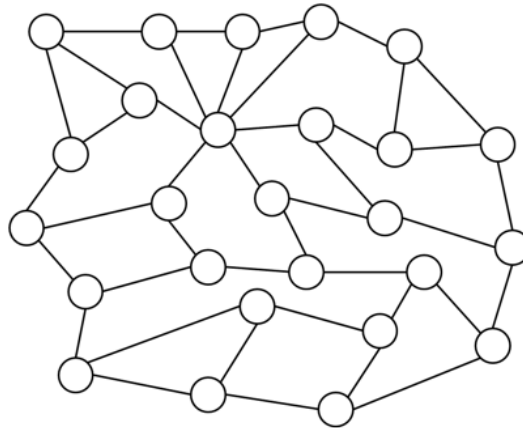


Abbildung 3.36: Beispiel für einen Graphen der Zusatzstudie

ne der Grafiken für einen bestimmten Zeitraum eingeblendet wurde. Dabei wurden drei verschieden lange Einblendungszeiten gewählt: 200 ms (also eine Länge, die der oberen Grenze der präattentiven Wahrnehmung entspricht), 600 ms und 1000 ms. Letztere beiden Werten basieren auf den Auswertungen der Aufzeichnungen der Hauptstudie durch den Eyetracker, bei denen eine sehr schnelle Beantwortung von Task 5 festgestellt wurde. Die Reihenfolge der Abbildungen mit und ohne einem höhergradigen Knoten, sowie die Reihenfolge der Einblendungszeiten, wurde zufällig permutiert, um zu verhindern, dass ein Lerneffekt eintritt, der das Ergebnis verfälschen könnte. An der Zusatzstudie nahmen zwölf Probanden teil, wobei es keine Überschneidungen zum Teilnehmerfeld bei der Hauptstudie gibt.

Die Ergebnisse der Auswertung liefert Abbildung 3.37. Hierbei muss berücksichtigt werden, dass die Frage, die die Probanden zu beantworten hatten, lautete: „Haben alle Knoten den gleichen Grad?“. Aus dieser Formulierung der Frage ergibt sich, dass die Teilnehmer bei Nicht-Erkennen eines höhergradigen Knotens automatisch mit „ja“ antworten. Dadurch werden alle die Antworten für die Auswertung hinfällig, bei denen tatsächlich ein Graph mit identischem Grad für alle Knoten vorliegt. Denn bei der Beantwortung kann nicht unterschieden werden, ob tatsächlich kein hochgradiger Knoten vorliegt, oder ob die Versuchsperson aufgrund mangelnder Konzentration oder Zeit nur die Standardantwort gegeben hat. Im Gegensatz dazu kann davon ausgegangen werden, dass bei Antwort „nein“ tatsächlich ein Knoten erkannt wurde, der einen höheren Grad aufweist. Damit bilden nur diese Graphen die Grundlage der Auswertung, die anderen Graphen sind jedoch trotzdem ein wichtiger Bestandteil der Versuchsanordnung, da mit ihrer Hilfe ein Lerneffekt effektiv unterbunden werden konnte.

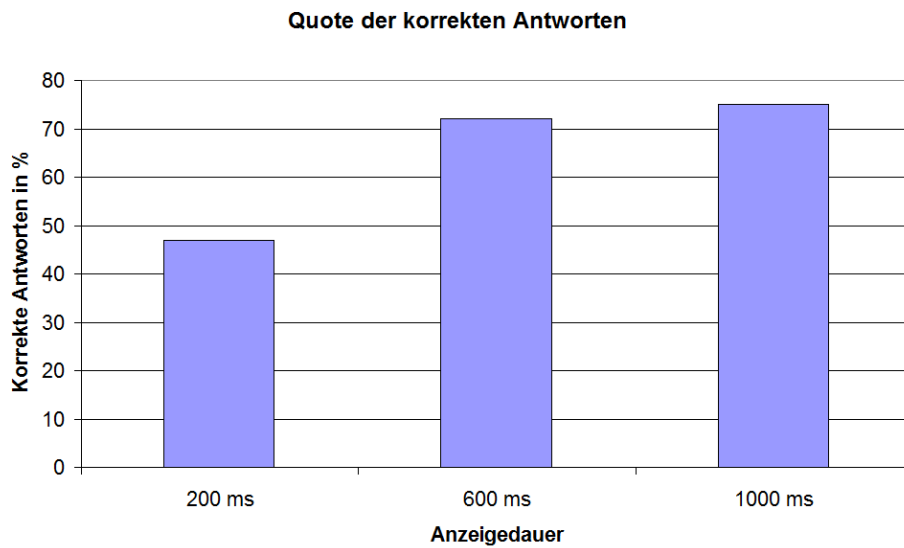


Abbildung 3.37: Ergebnis der Zusatzstudie

Die Quote der korrekten Antworten bei einer Einblendungszeit von 200 ms liegt bei dieser Studie bei nur 47%. Das bedeutet, die Probanden haben bei weniger als der Hälfte der ihnen gezeigten Graphen erkannt, dass ein Knoten mit einem signifikant höheren Grad vorliegt. Werden die gleichen Grafiken für 600 ms oder für 1000 ms eingeblendet, so steigen die Erfolgsquoten auf 72% beziehungsweise 75%. Obwohl zwischen den einzelnen Zeitstufen jeweils der gleiche Abstand von 400 ms liegt, ist also eine deutliche Verbesserung des Ergebnisses beim Übergang von 200 ms auf 600 ms zu erkennen, während die erneute Anhebung auf 1000 ms kaum noch zur Besserung der Quote beiträgt. Aus diesen Zahlen kann man erkennen, dass das Finden eines Kandidatenknotens mit einem hohen Grad, wie es in Task 5 gefordert war, zwar in sehr kurzer Zeit realisiert werden kann, allerdings haben sich die Anzeichen, dass es sich dabei um eine präattentive Wahrnehmung handeln könnte, nicht verfestigt. Hierbei spielt unter anderem auch die Lage des Knotens eine entscheidende Rolle, denn bei der Auswertung der einzelnen Abbildungen stellte sich zusätzlich heraus, dass ein hochgradiger Knoten in Randlage wesentlich schlechter erkannt wird als ein zentral gelegener Knoten. An dieser Stelle bleiben noch einige Fragen offen, die Bestandteil einer intensiveren zukünftigen Studie sein sollten.

4 Zusammenfassung und Ausblick

Abschließend soll im Folgenden ein kurzer Rückblick auf die Ergebnisse der Evaluation „Vergleich von Graphlayouts“ gegeben werden, bei der die wichtigsten Erkenntnisse nochmals aufgeführt werden. Im Anschluss daran folgen eine Beurteilung der Resultate und ein Ausblick auf Möglichkeiten zur Fortführung der Erforschung der Wahrnehmung verschiedener Graphlayouts.

4.1 Zusammenfassung

Diese Diplomarbeit soll einen Überblick darüber vermitteln, wie Graphen von menschlichen Betrachtern wahrgenommen werden, wo die Stärken und Schwächen liegen. Hierbei liegt der Schwerpunkt auf dem Vergleich von drei Graphlayoutverfahren: dem kräftebasierten, dem orthogonalen und dem hierarchischen Verfahren. Um die Zugänglichkeit dieser Layouts zu vergleichen, wurde eine Studie entwickelt, die, unter der Berücksichtigung der Ergebnisse einer vorgegangenen Machbarkeitsstudie, mit Hilfe eines Eyetrackingsystems das Blickverhalten der Probandengruppe erfasst. Diese Studie ist taskbasiert aufgebaut, das heißt die Zugänglichkeit der einzelnen Layouts wurde neben der Blicherfassung auch durch eine Performance - Messung (Korrektheit und durchschnittliche Bearbeitungszeit) überprüft.

Die Analyse gliedert sich nach den einzelnen Tasks auf. **Task 1** erwies sich als sehr einfache Aufgabe, die mit sehr hohen Korrektheitsquoten und einer sehr kurzen Bearbeitungszeit aufwarten kann. Die Besonderheit bei Task 1 liegt in der Unabhängigkeit vom verwendeten Layoutverfahren und von der Komplexität des Graphen. Denn zumindest bei den hier verwendeten Graphen lagen die Antwortzeiten zwischen den 10-Knoten und den 20-Knoten Graphen bei gerade einer Sekunde. Als Erklärung für diese Unabhängigkeit wurde festgestellt, dass diese Aufgabe keine adjazenten Beziehungen der Knoten beachtet, so dass die Layoutunterschiede hier nicht zum Tragen kommen. Die qualitative Analyse lieferte außerdem einen Hinweis

darüber, dass die Suchstrategie der Probanden sich nicht an strukturellen Besonderheiten der Layouts orientiert, sondern es findet ein „Scannen“ des Graphen von links oben nach rechts unten statt, was ebenfalls zur Layoutunabhängigkeit beiträgt.

Als sehr ergiebig erwies sich die Analyse von **Task 2**, dem Wegproblem. Hier konnten erste gravierende Unterschiede zwischen den getesteten Verfahren festgestellt werden. Während das kräftebasierte und das orthogonale Layout bei einer ersten Korrektheitsanalyse gleichauf liegen, weist das hierarchische Layout eine deutlich höhere Fehlerquote auf und auch die hohe time-out - Quote deutet auf große Zugänglichkeitsprobleme im Hinblick auf diese Aufgabenstellung hin. Dieser Trend bestätigte sich durch die Analyse der Bearbeitungszeiten, denn auch hier stieg der Aufwand bei hierarchischen Graphen mit steigender Komplexität überproportional stark an. In der anschließenden Analyse wurde untersucht, inwiefern die Layouts implizit das Konzept „kürzester Weg bei uniformen Pfadkosten“ unterstützen. Hierbei zeigten sich dann sehr deutliche Unterschiede zwischen dem kräftebasierten und dem orthogonalen Verfahren, denn während die Quote der korrekten Antworten nahezu gleich war, so wies das erstgenannte Layout eine hohe implizite Unterstützung kürzester Wege auf, während beim orthogonalen Graphen kaum kürzeste Wege genutzt wurden. Die Erklärung liegt in der mangelhaften Darstellung von adjazenten Beziehungen durch den orthogonalen Algorithmus. Hier würde eine Optimierung des orthogonalen Algorithmus Sinn machen, durch die benachbarte Knoten eher in räumlicher Nähe angeordnet werden, falls nötig auf Kosten der Kreuzungsminimierung, denn durch den orthogonalen Verlauf der Kanten zueinander sind Kreuzungen weniger problematisch als beim hierarchischen Layout. Bei diesem ist eine stärkere Optimierung in Hinsicht auf die Kreuzungsminimierung zu empfehlen, da durch geschwungene Kanten und zum Teil sehr spitze Winkel die Zugänglichkeit stark leidet. Bei der qualitativen Analyse wurde zudem festgestellt, dass die Betrachtungsweisen der Probanden bei den verschiedenen Layouts zum Teil völlig differieren: Während beim orthogonalen Layout die Knoten im Fokus der Fixationen stehen, so fixieren die Probanden beim hierarchischen Layout hauptsächlich die Kanten. Das kräftebasierte Layout liegt dabei zwischen diesen Extrema, hier werden eher die Knoten betrachtet, wenn es zu Kantenkreuzungen auf dem Weg kommt, dann rücken diese jedoch in den Mittelpunkt.

Task 3 und 4 lieferten aufgrund ihrer engen Verwandtschaft ähnliche Ergebnisse, was das kräftebasierte Layout anbelangt. Dabei wurde festgestellt, dass dieses Verfahren die Eigenschaft besitzt, die äußere Form eines Teilgraphen größtenteils zu erhalten und außerdem Cliques in einer sehr charakteristischen Form abzubilden. Das orthogonale Layout erreichte insbesondere bei Task 4 ein sehr schlechtes Ergebnis. Der Grund: Die teilweise Zerstörung der Wahrnehmung adjazenter Be-

ziehungen zwischen Knoten durch überlange Kanten. Beim Finden einer Clique, deren gesamtes Konzept auf Adjazenzen beruht, haben die Probanden daher erwartungsgemäß hohe Schwierigkeiten. Ähnlich schlecht schnitt auch das hierarchische Layout ab, die Gründe liegen hier ebenfalls in der Darstellung benachbarter Knoten. Durch deren Verteilung auf verschiedene Hierarchieschichten kommt es sowohl beim Teilgraph als auch besonders bei der Clique zu sehr langen Kanten. Durch die schlechte Kreuzungsminimierung entstehen daher sehr viele Kantenkreuzungen mit spitzen Winkeln, was das Erkennen der korrekten Lösung stark erschwert. Die qualitative Analyse bestätigte zudem die Ähnlichkeit bei beiden Tasks in der Suchstrategie der Versuchsteilnehmer. Diese suchten bei beiden Tasks layoutunabhängig nach bestimmten Formen, um somit Knotengruppen zu identifizieren, die eine mögliche Lösung darstellen. Nach dem Finden einer solchen Kandidatengruppe wurden dann die strukturellen Merkmale überprüft. Da jedoch das hierarchische und besonders das orthogonale Verfahren die Knoten einer korrekten Kandidatengruppe sehr weit voneinander entfernt abbilden können, führte diese Taktik hier nicht zum gewünschten Erfolg.

Task 5 wies wieder eine Ähnlichkeit zu Task 1 auf. Hier überraschte die quantitative Analyse mit der Erkenntnis, dass diese Aufgabenstellung offenbar weitestgehend unabhängig ist vom verwendeten Layout und der Komplexität der Graphen. Dies spiegeln sowohl die Ergebnisse der Korrektheitsanalyse als auch der Analyse der durchschnittlichen Bearbeitungszeiten wider. Die Erklärung hierfür lieferten die Heatmap-Analysen der verschiedenen Abbildungen. Bei allen Layoutverfahren wendeten die Probanden die gleiche Strategie an: Anstatt alle Knoten zu betrachten (was bei diesem Task zur korrekten Lösung eigentlich zwingend nötig ist), wurden jeweils nur drei bis fünf Kandidatenknoten fixiert, deren Umgebung ein dichteres Netz von Kanten aufweist. Diese wurden dann genauer überprüft und ein Knoten als Lösung ausgewählt, während alle anderen Knoten weitestgehend ignoriert wurden. Aus diesen Erkenntnissen heraus stellte sich nun die Frage, ob die Vorauswahl der Kandidatenknoten eine bewusste oder unbewusste Handlung der Probanden ist. Zur Klärung wurde eine Zusatzstudie durchgeführt, deren Ziel es war, Hinweise darauf zu finden, ob diese Wahrnehmung „präattentiv“ sein kann. Die Hinweise darauf verdichteten sich allerdings nicht, wobei man trotzdem feststellen kann, dass der größte Teil der Probanden innerhalb von einer Sekunde erkennt, ob ein höhergradiger Knoten vorliegt oder nicht.

4.2 Beurteilung und Ausblick

Diese Diplomarbeit stellt die (nach bestem Wissen) erste und bisher einzige wissenschaftliche Untersuchung dar, die verschiedene Graphlayoutverfahren, sowohl durch eine quantitative Analyse einer task-bezogenen Studie, als auch durch die Analyse des Blickverhaltens mittels eines Eyetrackingsystems, gegenüberstellt. Um die entsprechenden Erkenntnisse zu erlangen war ein hoher Aufwand an Labor- und Analysezeit notwendig, bei dem (mangels Existenz) kaum auf frühere Forschungsergebnisse zurückgegriffen werden konnte. Die wenigen Studien, bei denen Eyetrackingsysteme zur Anwendung kamen, beschäftigen sich entweder mit einem speziellen Layout, bei dem verschiedene Graphen zum Einsatz kommen (z.B. [Hua05]), oder sie betrachten hauptsächlich die allgemeinen Ästhetikkriterien. Auch die Mehrzahl der rein empirischen Evaluationen beschäftigt sich weniger mit Graphlayoutverfahren, als mit der Beurteilung von ästhetischen Kriterien (z.B. [Pur02], [Pur97], [War02]). Besonders im Kontext von UML-Diagrammen findet man zusätzlich einige Eyetracking-Studien (z.B. [Yus07]), die allerdings ebenfalls kaum Vergleiche verschiedener Darstellungsarten liefern.

Durch die aussagekräftigen Ergebnisse ist der Grundstein für eine weitere Forschung in dieser Thematik gelegt. Desweiteren können die verwendeten Layoutalgorithmen entsprechend der taskbasierten Ergebnisse optimiert werden. Dabei ist darauf zu achten, dass die anvisierten Verbesserungen alle task-spezifisch sind, d.h. die Optimierungen sind immer in dem Kontext zu sehen, für den der jeweilige Graph gezeichnet wird. Das kräftebasierte Verfahren erwies sich bei dieser Studie als ein sehr gutes Allroundverfahren, das in keiner Disziplin echte Schwächen aufweist. Demgegenüber stehen das orthogonale und das hierarchische Layout, die bei einzelnen Aufgaben nur unterdurchschnittlich abgeschlossen haben. Dabei muss allerdings beachtet werden, dass die verwendeten Graphen von Hand entworfen wurden und in keinen Kontext eingebettet sind. Es ist durchaus vorzustellen, dass bei Graphen, deren Quelldaten eine hierarchische Ordnung aufweisen, das hierarchische Verfahren bei entsprechenden Aufgaben seine Stärken besser zur Geltung bringen könnte. Dies kann Gegenstand einer zukünftigen kontextbezogenen Studie sein, die auf den Erkenntnissen dieser Evaluation beruht.

Es muss ferner angemerkt werden, dass alle Ergebnisse auf Graphen mit maximal 20 Knoten beruhen. Im Hinblick auf die Annahme, dass Task 1 und Task 5 wesentlich resistenter gegenüber der Erhöhung der Graphkomplexität sind, als es bei den anderen Aufgaben der Fall ist, wäre es interessant zu sehen, ob eine signifikante Erhöhung der Knotenzahl und/oder des Durchschnittsgrades diese Beobachtungen

bestätigen. Auch bei den anderen Tasks wären Studien mit komplexeren Graphen vermutlich sehr ergiebig und ein lohnender Forschungsgegenstand. Wie bei der Vorstellung des verwendeten Layoutprogramms DGD bereits erwähnt, nutzten die hier zur Verwendung gekommenen Graphen nur einen Bruchteil der Möglichkeiten dieses Tools aus, das seine Stärke insbesondere in der Darstellung dynamischer Graphen hat. Auch hier würde eine Analyse von über die Zeit veränderlichen Graphen in verschiedenen Layouts mittels Eyetracking und task-spezifischer Analyse wissenschaftliches Neuland betreten. Der Weg dorthin wurde durch diese Arbeit bereits angelegt.

Literaturverzeichnis

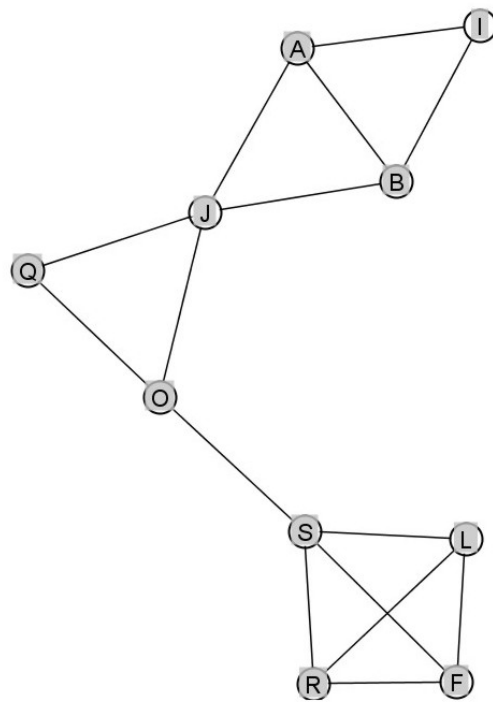
- [Bas99] BASTERT, Oliver und MATUSZEWSKI, Christian: Layered Drawings of Digraphs, in: *Drawing Graphs*, S. 87–120
- [Bat89] BATTISTA, Giuseppe Di; TAMASSIA, Roberto und TOLLIS, Ioannis G.: Area Requirement and Symmetry Display in Drawing Graphs, in: *Symposium on Computational Geometry*, S. 51–60
- [Bat99] BATTISTA, Giuseppe Di; DIDIMO, Walter; PATRIGNANI, Maurizio und PIZZONIA, Maurizio: Orthogonal and Quasi-upward Drawings with Vertices of Prescribed Size, in: *Graph Drawing*, S. 297–310
- [Bie94] BIEDL, Therese C. und KANT, Goos: A Better Heuristic for Orthogonal Graph Drawings, in: *ESA*, S. 24–35
- [Bra99] BRANDES, Ulrik: Drawing on Physical Analogies, in: *Drawing Graphs*, S. 71–86
- [Bra02] BRANDES, Ulrik; EIGLSPERGER, Markus; KAUFMANN, Michael und WAGNER, Dorothea: Sketch-Driven Orthogonal Graph Drawing, in: *Graph Drawing*, S. 1–11
- [Car80] CARPANO, M.: Automatic display of hierarchized graphs for computer aided decision analysis (1980), Bd. SE-12(4):S. 538–546
- [Cof72] COFFMAN, Edward G. und GRAHAM, Ronald L.: Optimal Scheduling for Two-Processor Systems. *Acta Inf.* (1972), Bd. 1:S. 200–213
- [del] <http://del.icio.us/>, zuletzt besucht am 15. Februar 2008
- [Die06] DIESTEL, R.: *Graphentheorie*, Bd. 3., neu bearb. und erw. A., Springer, Berlin (2006)

- [Duc02] DUCHOWSKI, A. T.: *Eye Tracking Methodology: Theory and Practice* (2002)
- [Ead84] EADES, Peter und MCKAY, Brendan D.: An Algorithm for Generating Subsets of Fixed Size With a Strong Minimal Change Property. *Inf. Process. Lett.* (1984), Bd. 19(3):S. 131–133
- [Ead90] EADES, Peter und SUGIYAMA, Kozo: How to draw a directed graph. *J. Inf. Process.* (1990), Bd. 13(4):S. 424–437
- [Ead94] EADES, Peter und WORMALD, Nicholas C.: Edge Crossings in Drawings of Bipartite Graphs. *Algorithmica* (1994), Bd. 11(4):S. 379–403
- [Eig99] EIGLSPERGER, Markus; FEKETE, Sándor P. und KLAU, Gunnar W.: Orthogonal Graph Drawing, in: *Drawing Graphs*, S. 121–171
- [For90] FORMANN, Michael; HAGERUP, Torben; HARALAMBIDES, James; KAUFMANN, Michael; LEIGHTON, Frank Thomson; SYMVONIS, Antonios; WELZL, Emo und WOEGINGER, Gerhard J.: Drawing Graphs in the Plane with High Resolution, in: *FOCS*, S. 86–95
- [Föß95] FÖSSMEIER, Ulrich und KAUFMANN, Michael: Drawing High Degree Graphs with Low Bend Numbers, in: *Graph Drawing*, S. 254–266
- [Fri94] FRICK, Arne; LUDWIG, Andreas und MEHLDAU, Heiko: A Fast Adaptive Layout Algorithm for Undirected Graphs, in: Roberto Tamassia und Ioannis G. Tollis (Herausgeber) *Proc. DIMACS Int. Work. Graph Drawing, GD*, 894, Springer-Verlag, Berlin, Germany, S. 388–403, URL citeseer.ist.psu.edu/frick94fast.html
- [Fru91] FRUCHTERMAN, Thomas M. J. und REINGOLD, Edward M.: Graph Drawing by Force-directed Placement. *Softw., Pract. Exper.* (1991), Bd. 21(11):S. 1129–1164
- [Gan93] GANSNER, Emden R.; KOUTSOFIOS, Eleftherios; NORTH, Stephen C. und VO, Kiem-Phong: A Technique for Drawing Directed Graphs. *IEEE Trans. Software Eng.* (1993), Bd. 19(3):S. 214–230
- [gdc] <http://www.cs.usyd.edu.au/~visual/gd2007/>, zuletzt besucht am 15. Februar 2008

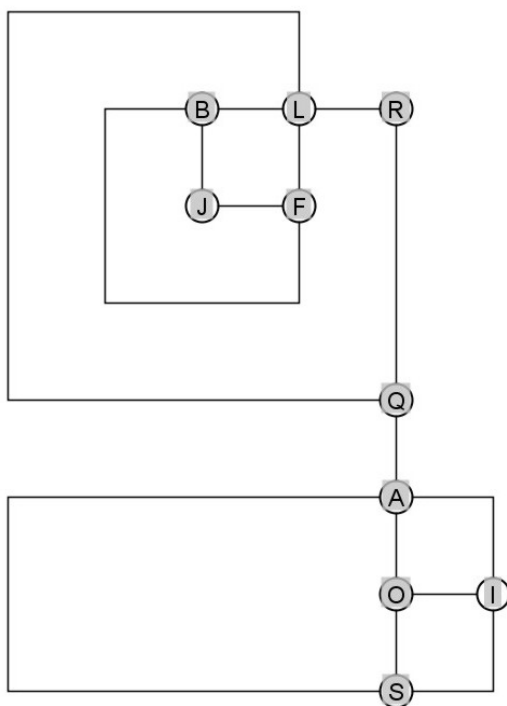
- [Gre87] GREENGARD, Leslie Frederick: The rapid evaluation of potential fields in particle systems, Yale University, New Haven, CT, USA
- [Hea06] HEALEY, C.G.: Perception in Visualization, <http://www.csc.ncsu.edu/faculty/healey/PP/> (2006), URL <http://www.csc.ncsu.edu/faculty/healey/PP/>
- [Hua05] HUANG, Weidong und EADES, Peter: How People Read Graphs, in: *AP-VIS*, S. 51–58
- [inf] <http://www.infovis.org/>, zuletzt besucht am 15. Februar 2008
- [Kam88] KAMADA, Tomihisa und KAWAI, Satoru: Advanced graphics for visualization of shielding relations. *Computer Vision, Graphics, and Image Processing* (1988), Bd. 43(3):S. 294–312
- [Lev91] LEVEN, Wilfried: *Blickverhalten von Konsumenten. Grundlagen, Messung und Anwendung in der Werbeforschung*, Physica-Verlag Heidelberg (1991)
- [Mal94] MALITZ, Seth M. und PAPA KOSTAS, Achilleas: On the Angular Resolution of Planar Graphs. *SIAM J. Discrete Math.* (1994), Bd. 7(2):S. 172–183
- [Mat89] MATSUOKA, Satoshi; KAMADA, Tomihisa und KAWAI, Satoru: Asymptotic evaluation of window visibility. *Inf. Process. Lett.* (1989), Bd. 32(3):S. 119–126
- [Mis95] MISUE, Kazuo; EADES, Peter; LAI, Wei und SUGIYAMA, Kozo: Layout Adjustment and the Mental Map. *J. Vis. Lang. Comput.* (1995), Bd. 6(2):S. 183–210
- [Neu99] NEUMANN, S.: Praktische Vorgehensweise bei der Konzeption, Durchführung und Auswertung von Augenbewegungsmessungen bei experimentellen kartographischen Untersuchungen (1999)
- [Pur97] PURCHASE, Helen C.: Which Aesthetic has the Greatest Effect on Human Understanding?, in: *Graph Drawing*, S. 248–261
- [Pur02] PURCHASE, Helen C.; CARRINGTON, David A. und ALLDER, Jo-Anne: Empirical Evaluation of Aesthetics-based Graph Layout. *Empirical Software Engineering* (2002), Bd. 7(3):S. 233–255

- [San96] SANDER, Georg: *Visualisierungstechniken für den Compilerbau*, Dissertation, Universität des Saarlandes (1996)
- [Ser07] SERVATIUS, Kerstin: Einführung in die Blickbewegungsregistrierung (2007)
- [stu] <http://www.studivz.net>, zuletzt besucht am 15. Februar 2008
- [Sug81] SUGIYAMA, K.; TAGAWA, S. und TODA, M.: Methods for Visual Understanding of Hierarchical Systems. *IEEE Trans. Syst. Man Cybern.* (1981), Bd. SMC-11(2):S. 109–125
- [Tam87] TAMASSIA, Roberto: On Embedding a Graph in the Grid with the Minimum Number of Bends. *SIAM J. Comput.* (1987), Bd. 16(3):S. 421–444
- [Tob05] TOBII: *Tobii User Manual: Tobii Eye Tracker, ClearView analysis software*, Tobii Technology AB Schweden (2005)
- [War77] WARFIELD, John: Crossing Theory and Hierarchy Mapping (1977), Bd. SMC-7(7):S. 505–523
- [War02] WARE, Colin; PURCHASE, Helen C.; COLPOYS, Linda und MCGILL, Matthew: Cognitive measurements of graph aesthetics. *Information Visualization* (2002), Bd. 1(2):S. 103–110
- [Weg07] WEGENER, Ingo: *Komplexitätstheorie. Grenzen der Effizienz von Algorithmen*, Springer, Berlin (2007)
- [wik] <http://de.wikipedia.org>, zuletzt besucht am 15. Februar 2008
- [Yus07] YUSUF, Shehnaaz; KAGDI, Huzefa H. und MALETIC, Jonathan I.: Assessing the Comprehension of UML Class Diagrams via Eye Tracking, in: *ICPC*, S. 113–122

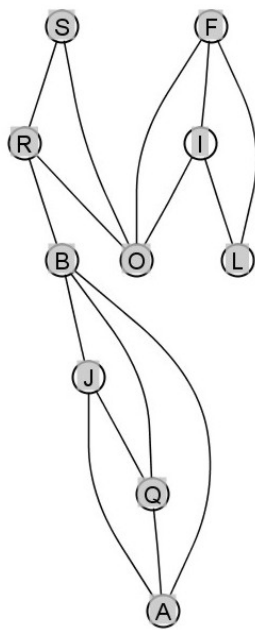
Anhang



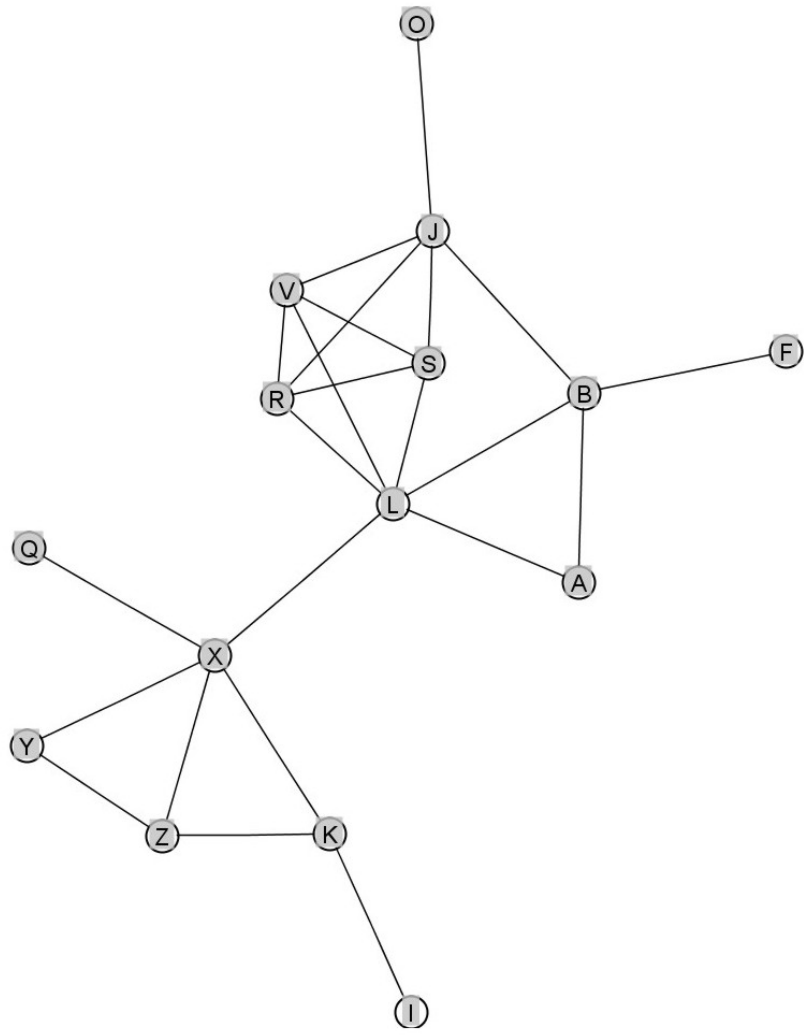
10 Knoten, kräftebasiert



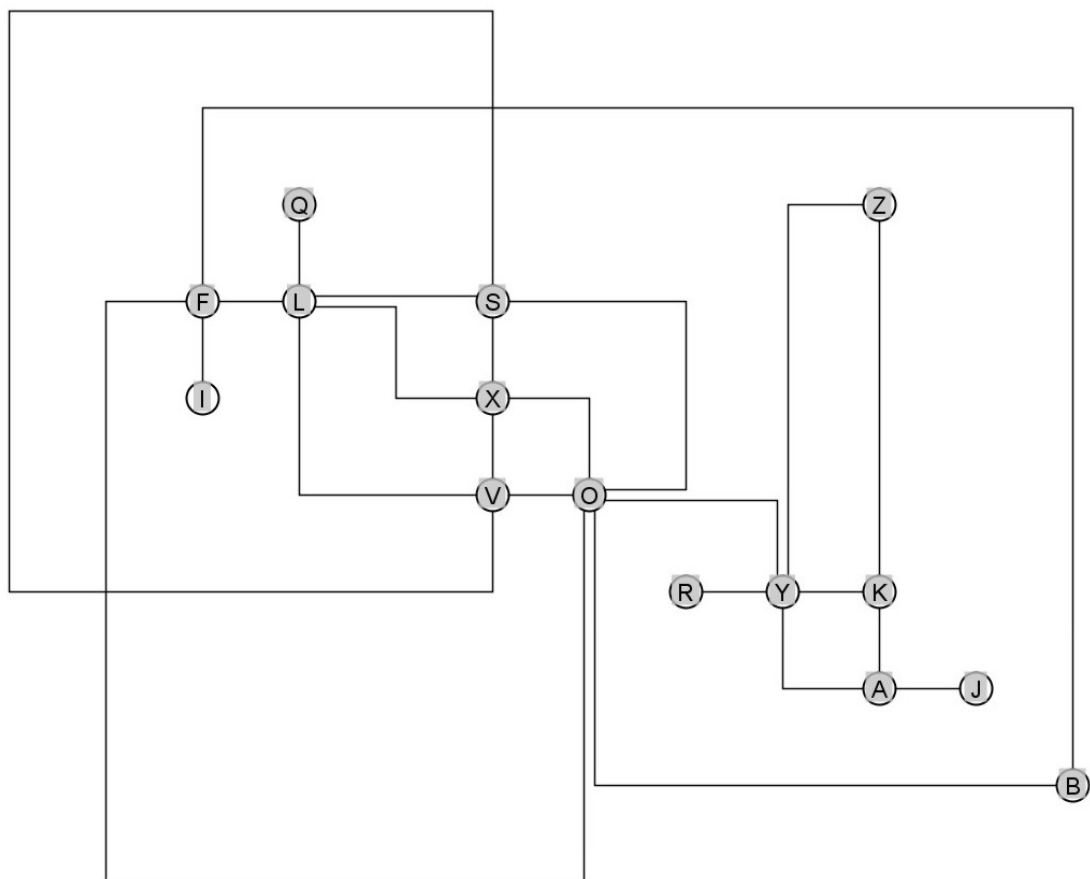
10 Knoten, orthogonal



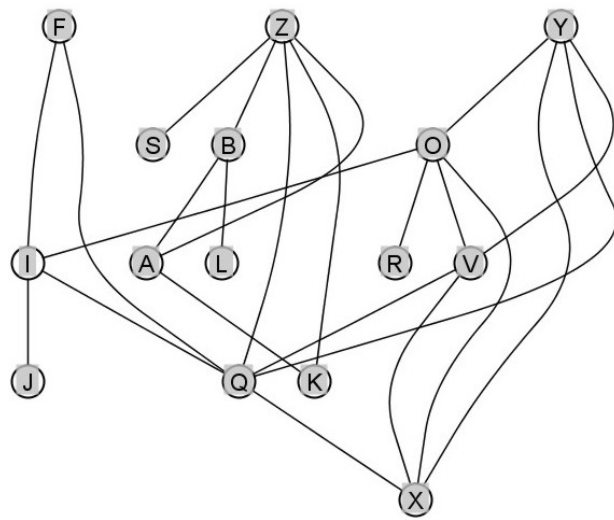
10 Knoten, hierarchisch



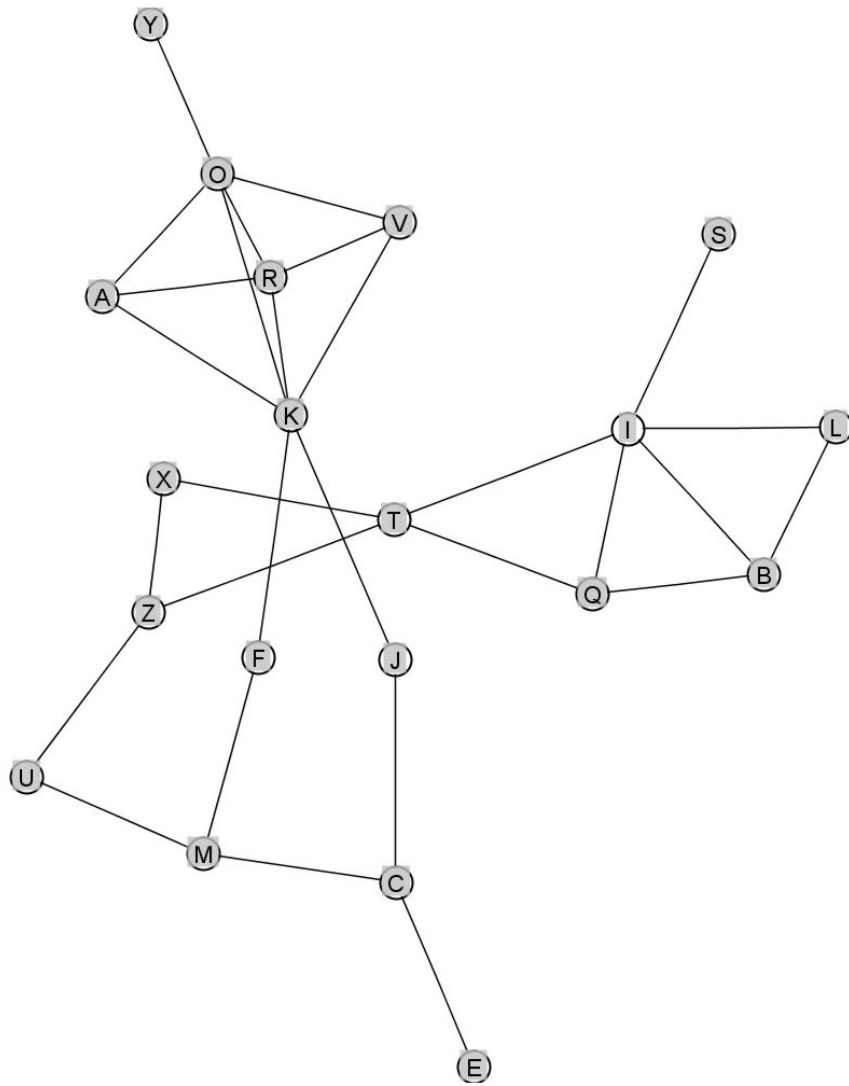
15 Knoten, kräftebasiert



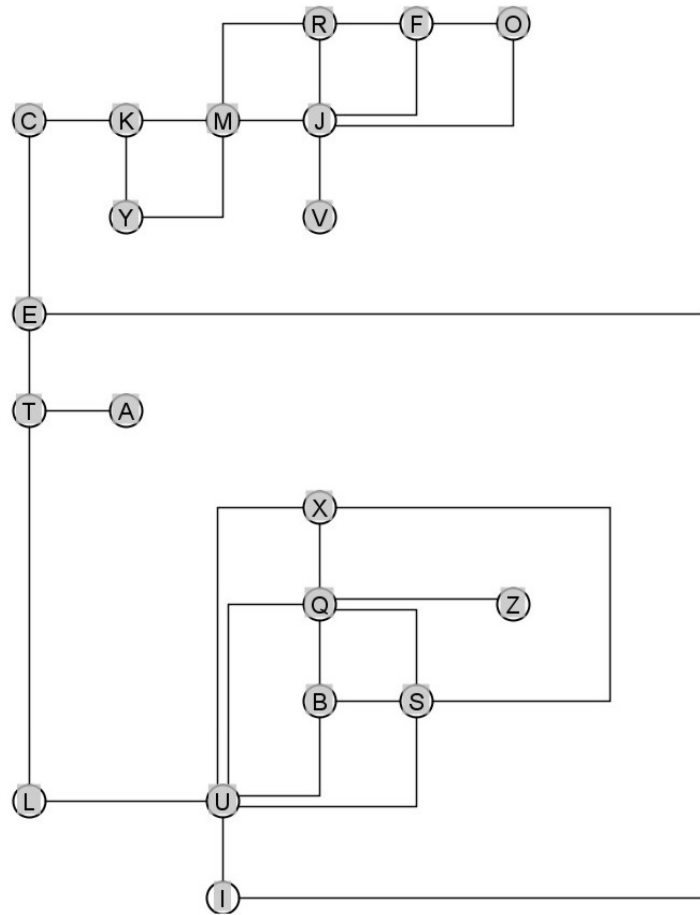
15 Knoten, orthogonal



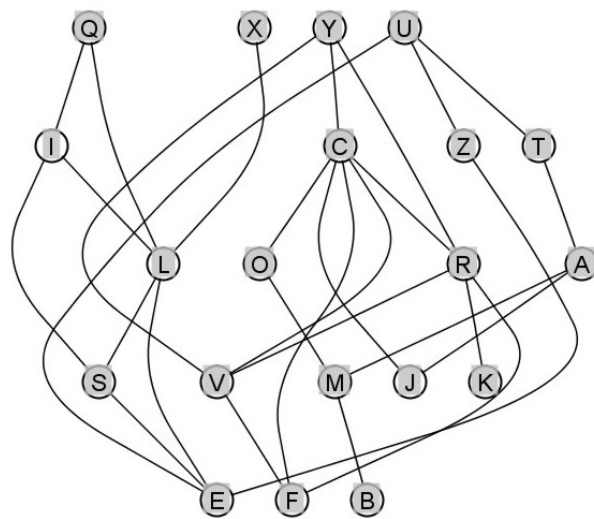
15 Knoten, hierarchisch



20 Knoten, kräftebasiert



20 Knoten, orthogonal



20 Knoten, hierarchisch