Automatisiertes Zeichnen von Diagrammen*

Petra Mutzel

Max-Planck-Institut f. Informatik
Im Stadtwald, D-66123 Saarbrücken

Ein sehr anwendungsnahes Forschungsgebiet, bei dem viele NP-schwere Optimierungsprobleme auftauchen ist das automatisierte Zeichnen von Diagrammen bzw. Graphen. Hierbei sind Objekte zusammen mit ihren Relationen gegeben, deren Strukturen möglichst übersichtlich bildlich dargestellt werden sollen. So sind zum Beispiel in chemischen Reaktionsdiagrammen die chemischen Elemente zusammen mit ihren möglichen Reaktionen gegeben. In einem Bild veranschaulicht lassen sich die Wechselwirkungen untereinander viel leichter erkennen.

Das Sprichwort "Bilder sagen mehr als tausend Worte" trifft auch für viele andere Anwendungsbereiche zu, wie zum Beispiel der Veranschaulichung von Relationen in Datenbanken, Animation von Programmen, Zeichnen von PERT-, ER- oder Flußdiagrammen. Anwendungen tauchen im wirtschaftswissenschaftlichen Umfeld genauso auf wie im chemischen, physikalischen oder dem informationstechnischen Bereich.

In den meisten Fällen werden diese Diagramme mit der Hand angefertigt. Zunehmend verläßt man das traditionelle Handwerkszeug "Zirkel und Lineal" und verwendet Computerprogramme, die das Zeichnen selbst sehr erleichtern. Selbst das computerunterstützte Zeichnen von Diagrammen bleibt jedoch eine zeitraubende Angelegenheit.

Mit wachsender Größe der zu zeichnenden Diagramme kommt sehr schnell die Überschaubarkeit abhanden, so daß es immer seltener gelingt, diese Diagramme "übersichtlich" zu zeichnen. Doch selbst bei relativ kleinen Diagrammen ist dies schon oft der Fall.

Das in Abbildung 1 gezeigte Reaktionsflußdiagramm hat uns Holger Beck, ein Astrophysiker der Technischen Universität Berlin zur Verfügung gestellt, der sich mit der Chemie in den Hüllen von Sternen beschäftigt. In der Praxis hat sich gezeigt, daß es schwierig ist, selbst für diese relativ kleinen Diagramme, eine übersichtliche, möglichst überkreuzungsarme Darstellung zu finden.

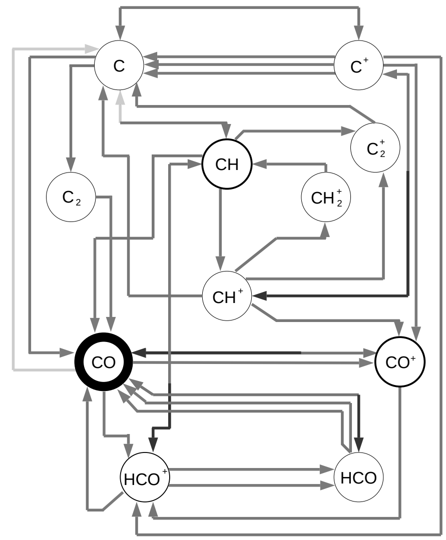

Abb. 1: computerunterstützt gezeichnetes chemisches Reaktionsflußdiagramm.

Ein Programm, das automatisch die strukturelle Information auswertet und das Diagramm "schön" und "ästhetisch" zeichnet, wäre hier wünschenswert.

Historisch aus dem CAD (Computer Aided Design) entstanden, hat sich das Gebiet des Automatisierten Zeichnens von Diagrammen und Graphen inzwischen zu einem selbständigen und lebendigen Forschungsgebiet entwickelt. Seit 1993 gibt es eine speziell zu diesem Themengebiet jährliche internationale Konferenz, deren Teilnehmerzahl stetig steigt, wobei ein hoher Prozentsatz der Teilnehmer (ca. 20% im Jahr 1994) aus der Industrie kommt.

Die Methoden zum automatisierten Zeichnen allgemeiner Diagramme kann man in drei Klassen einteilen. Ein Verfahren, das auf einem physikalischen Kräftemodell beruht, eine Methode, die auf einer Zerlegung des Diagrammes in mehrere Schichten beruht und ein Verfahren, das die Vorteile "planarer Graphen" ausnützt. Das letztere Verfahren zusammen mit den dabei auftretenden Problemen werden wir hier detailliert erläutern. Doch zunächst wenden wir uns der Frage zu, inwieweit bereits Systeme zum automatisierten Zeichnen zur Verfügung stehen.

Die bisherigen Systeme zum automatischen Zeichnen von Diagrammen basieren meistens auf einem der beiden erstgenannten Verfahren, da diese sehr einfach zu verstehen und zu im-

* geschrieben am 30.03.95 für das Jahrbuch 1995 der Max-Planck-Gesellschaft

plementieren sind. Die Qualität der weltweit erreichten Zeichnungen gibt noch viele Mängel zu erkennen. Es wird noch viel Forschung investiert werden müssen, bis wirklich marktfähige Programme zum automatisierten Zeichnen entwickelt worden sind.

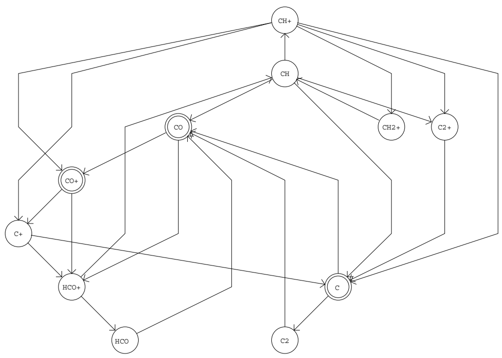

Abb. 2: Zeichnung mit dem Schichtenverfahren.

Abbildung 2 zeigt eine automatisch generierte Zeichnung des in Abbildung 1 gezeigten chemischen Reaktionsdiagramms. Die Zeichnung wurde mit dem in Passau von Michael Himsolt und Franz-Josef Brandenburg entwickelten "GraphEd" System gezeichnet. Als Zeichenverfahren wurde das Schichtenmodell gewählt. Da "GraphEd" keine Mehrfachkanten erfassen kann, wurde je nur eine Linie zwischen zwei Objekten gezeichnet. Eine andere Möglichkeit liegt darin, das physikalische Kräftemodell zu wählen. Das Ergebnis ist in Abbildung 3 zu sehen.

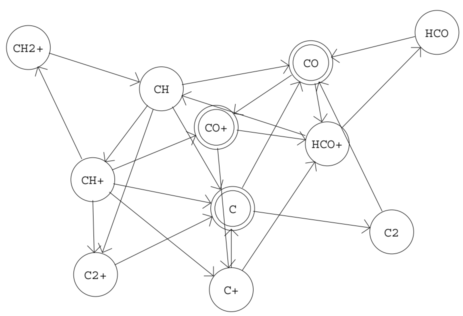

Abb. 3: Zeichnung mit dem physikalischen Modell.

In Kooperation mit Franz-Josef Brandenburg, Michael Himsolt (Universität Passau), Michael Jünger (Universität zu Köln) und Stefan Näher entwickeln wir im Rahmen des Schwerpunktprogramms "Effiziente Algorithmen" ein System, das alle genannten Verfahren zum Zeichnen von Graphen zur Verfügung stellen wird und zusätzlich viele Zeichenverfahren für spezielle Diagramme enthalten wird. Darüber hinaus wird es für den Benutzer erweiterbar sein, so daß jeder ohne weiteres seine eigenen Verfahren hinzufügen kann. Dies wird erleichtert durch die vorhandene Schnittstelle zu LEDA, die am Institut zusammen mit Stefan Näher entwickelte Software-Bibliothek effizienter Datenstrukturen und Algorithmen.

Das in Passau entwickelte Vorgängersystem "GraphEd" enthält bereits teilweise die geforderten Funktionen und wird bereits vielfältig in Wirtschaft, Industrie und Universität eingesetzt.

Mit dem neu entwickelten System wird es möglich sein, die Zeichnungen der verschiedenen Zeichenverfahren innerhalb desselben Systems zu vergleichen. Um Aussagen über die Qualität der Zeichnungen machen zu können, brauchen wir Kriterien für eine "schöne" Zeichnung. Solche objektive Kriterien sind zum Beispiel, eine ausgewogene Verteilung der Objekte und Linien in der Zeichnung, eine kleine Anzahl von Überkreuzungen der Linien oder eine kleine Anzahl von Knicken in den Linien.

Außerdem wäre es wichtig, eine Bibliothek typischer Diagramme zu haben, anhand derer die verschiedenen Zeichenverfahren ausgewertet werden können. Eine solche Diagramm-Bibliothek wollen wir in Kooperation mit Stephen North (AT&T Bell Laboratories, New Jersey), Michael Himsolt (Universität Passau) und Thomas Lange (Universität zu Köln) erstellen.

Wie schon angekündigt, werden wir jetzt das Zeichenverfahren, das die Vorteile der "planaren Graphen" ausnützt, zusammen mit den dabei auftretenden mathematischen Problemen genauer erläutern.

Dazu gehen wir auf eine abstraktere Form von Diagrammen über. Diagramme können als Graphen modelliert werden. Ein Graph ist eine Menge von Knoten (Objekten bzw. Punkten) und Kanten (Linien), wobei eine Kante jeweils zwei Knoten verbindet. Die Kanten können gerichtet (wie in Abbildung 1) oder auch ungerichtet sein. Nun steht uns die ganze wunderschöne mathematische Theorie der Graphen, insbesondere die der planaren Graphen, zur Verfügung.

Läßt sich der Graph auf solche Weise in die Ebene zeichnen, daß sich keine der Linien überkreuzen, so spricht man von einem *planaren Graphen*. Da die Theorie planarer Graphen in unserer Forschung einen wesentlichen Anteil hat, werden wir eine kurzen Einblick in diese Theorie geben.

Die erste Referenz auf planare Graphen gibt es 1736 mit dem berühmten Satz von Leonhard Euler, der besagt, daß die Anzahl der Kanten in einem planaren Graphen höchstens die dreifache Anzahl der Knoten minus sechs sein kann. Doch die Theorie der planaren Graphen wurde erst 1930 von Kazimierz Kuratowski begründet. Er konnte nämlich zeigen, daß ein Graph genau dann planar ist, wenn er weder eine Unterteilung vom vollständigen Graphen auf 5 Knoten (d.h. jedes Knotenpaar ist miteinander durch eine Kante verbunden) noch eine Unterteilung vom vollständigen zweigeteilten Graphen auf 3×3 Knoten (d.h. jeder der drei Knoten in der einen Menge ist mit jedem der anderen Menge mit einer Kante verbunden) enthält. Bei einer Unterteilung eines Graphen ist es erlaubt, beliebige zusätzliche Knoten auf den ursprünglichen Kanten anzubringen. Abbildung 4 zeigt jeweils eine Unterteilung der beiden verbotenen Untergraphen.

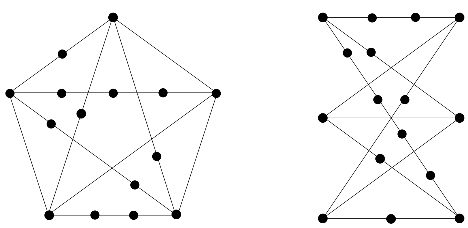

Abb. 4: Unterteilungen der verbotenen Untergraphen.

Beide Graphen spielten schon vorher in verschiedenen mathematischen Rätseln eine zentrale Rolle. Wer kennt nicht das Rätsel, das Dudeney 1917 in seinem berühmten Rätselbuch in eine moderne Form brachte:

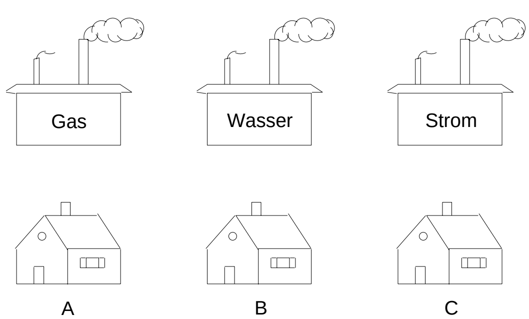

Abb. 5: Das Rätsel Gas–Wasser–Strom.

Jedes der Häuser A, B und C soll Wasser, Gas und Strom bekommen. Können die Leitungen so gelegt werden, ohne daß sie sich überkreuzen? Eine etwas kniffligere Version entsteht, wenn man zuläßt, daß die Häuser jeweils an beliebige Orte plaziert werden können.

Die Tatsache, daß sich die verbotenen Untergraphen nicht überkreuzungsfrei in die Ebene zeichnen lassen, war wohl schon länger bekannt. Die Rückrichtung jedoch, daß dies die beiden einzigen verbotenen Untergraphen sind, ist doch sehr überraschend. Kazimierz Kuratowski hat durch diese Arbeit die topologische Graphentheorie wiederbelebt, die seit der Gründung durch Leonhard Euler fast 200 Jahre lang keinerlei Beachtung fand.

Für die Klasse planarer Graphen existieren einige Zeichenverfahren, die viele der objektiven Kriterien für eine "schöne" Zeichnung erfüllen. Erlaubt man keine Knicke der Kanten, dann kann man garantieren, daß sich der Platzverbrauch einer überkreuzungsfreien Zeichnung höchstens quadratisch zur Anzahl der Knoten verhält. Ein anderes Verfahren garantiert, daß fast jede Fläche, die durch die Linien des in die Ebene gezeichneten Graphen begrenzt wird, konvex gezeichnet wird. Ist die Anzahl der benachbarten Knoten eines Knotens kleiner gleich vier, so existiert ein Verfahren, das nur horizontale und vertikale Linien zeichnet und dabei (bei gegebener planarer Einbettung) die totale Anzahl der Knicke in den Linien minimiert.

Um diese speziellen Verfahren auch für allgemeine (nicht-planare) Graphen ausnützen zu können, versucht man, eine minimale Anzahl von Kanten aus dem gegebenen Graphen zu entfernen, so daß der resultierende Graph planar ist.

Nun stehen zwei Möglichkeiten zur Auswahl. Eine Möglichkeit ist, den planaren Graphen mit den oben genannten Methoden zu zeichnen und danach die entfernten Kanten in die Zeichnung wiedereinzufügen. Dies kann automatisch oder interaktiv geschehen. In den meisten Fällen sind die in der Praxis auftretenden Graphen "fast" planar, d.h. die Anzahl der zu entfernenden Kanten ist sehr klein (zwischen 1 und 7).

Eine zweite Möglichkeit besteht darin, vor dem Zeichnen die entfernten Kanten nun wieder in den Graphen einzufügen, so daß die Anzahl der Überkreuzungen möglichst gering wird. Jede auftretende Überkreuzung von Kanten wird danach durch künstliche Knoten ersetzt, d.h. die beiden kreuzenden Kanten werden durch einen neuen Knoten und vier Kanten ersetzt. Der so künstlich planarisierte Graph kann nun mit den speziellen Verfahren für planare Graphen gezeichnet werden. Diese Methode, nicht-planare Graphen zu zeichnen, beinhaltet viele Schritte, die im folgenden detaillierter erläutert werden.

Planarisierung. Zunächst muß eine kleinste Kantenmenge bestimmt werden, deren Entfernung zu einem planaren Graphen führt. Dieses Problem ist äquivalent mit dem Problem in einem gegebenen nicht-planaren Graphen einen größten planaren Untergraphen zu bestimmen. Das größte planare Untergraphenproblem ist seit vielen Jahren als schweres Problem bekannt, d.h. es ist kein effizientes Verfahren zur Bestimmung größter planarer Untergraphen bekannt und es wird auch bezweifelt, daß ein solches existiert (im komplexitätstheoretischen Sinn ist das Problem NP-schwer).

Hier ist es uns in Zusammenarbeit mit Michael Jünger (Universität zu Köln) gelungen, ein Verfahren zu entwickeln, das für Graphen moderater Größenordnung mit hoher Wahrscheinlichkeit tatsächlich einen größten planaren Untergraphen bestimmt. Die Idee dieses Verfahrens sei hier nur sehr kurz erläutert.

Zunächst wird der Menge aller planaren Untergraphen des gegebenen Graphen die Menge der Ecken eines hochdimensionalen Polyeders zugeordnet. In mathematischer Feinarbeit werden dann die strukturellen Eigenschaften dieses Polyeders untersucht. Die Kenntnisse über die Struktur des Polyeders werden algorithmisch in sogenannten Schnittebenentechniken ausgenutzt.

So konnten wir nachweisen, daß die verbotenen Untergraphen in Kazimiersz Kuratowski's Arbeiten eine wesentliche Rolle in diesem Polyeder spielen. Aber auch die Euler'sche Formel für planare Graphen findet sich in dem Polyeder wieder. Zugleich fanden wir einige Formeln, die die Konzepte von Euler und Kuratowski verallgemeinern.

Schnittebenenverfahren werden seit wenigen Jahren, zum Teil sehr erfolgreich, in der Optimierung, so zum Beispiel für das Handlungsreisendenproblem, eingesetzt.

Mit unserem in Kooperation mit Stefan Thienel und Michael Jünger (Universität zu Köln) implementierten Verfahren gelang es uns, die größten planaren Untergraphen von vielen in der Literatur vorgegebenen Instanzen zum ersten Mal zu bestimmen.

Topologische Einbettung. Um den erhaltenen planaren Graphen nun auch überkreuzungsfrei zeichnen zu können, muß zunächst eine topologische Einbettung erstellt werden. Eine topologische Einbettung ist eine Liste aller Knoten zusammen mit einer geordneten Liste der Nachbarknoten per Knoten, so daß eine ebene (kreuzungsfreie) Darstellung des Graphen mit genau dieser Ordnung existiert.

Das Problem, eine topologische Einbettung eines planaren Graphen zu erstellen, ist verhältnismäßig leicht. In der Literatur werden im wesentlichen zwei verschiedene Verfahren genannt, deren Laufzeit proportional zur Anzahl der Knoten, also in schnellstmöglicher Zeit, ist. Einer davon, derjenige, der auf dem Planaritätstest von Hopcroft und Tarjan beruht, wird jedoch nur ungenügend skizziert. Da die eigentliche Problematik sehr komplex und nicht dargestellt ist, kommt es immer wieder zu Mißverständnissen, die falsche Implementierungen dieses Algorithmus nach sich ziehen. In verschiedenen Arbeiten zeigen wir (Kurt Mehlhorn und Petra Mutzel) die Problematik auf und geben detaillierte Verfahren an, die eine topologische Einbettung eines gegebenen planaren Graphen mit Hilfe des Planaritätstests von Hopcroft und Tarjan zu berechnen.

Implementierungen beider Verfahren zeigen, daß wir nun zum Beispiel die topologische Einbettung von Graphen mit 30000 Knoten und 64000 Kanten in weniger als 50 Sekunden auf einem PC berechnen können.

Augmentierung. Der durch Planarisierung erhaltene Graph soll nun gezeichnet werden. Viele Zeichenverfahren erfordern zweizusammenhängende planare Graphen als Eingabe, d.h. zwischen jedem Knotenpaar müssen mindestens zwei knotendisjunkte Wege existieren. Ist der ursprünglich gegebene, nicht-planare Graph zweizusammenhängend, dann könnte man anstatt der beschriebenen Planarisierung versuchen, gleich einen größten zweizusammenhängenden planaren Untergraphen zu bestimmen, der jedoch alle Knoten enthalten sollte. Dieses Problem gehört zu den "schweren" Problemen (im Sinne der Komplexitätstheorie). Selbst wenn der gegebene Graph zweizusammenhängend ist, ist die Existenz eines spannenden zweizusammenhängenden Untergraphen nicht garantiert. Es bleibt die Möglichkeit, neue Kanten zu dem gegebenen Graphen hinzuzufügen, so daß der Graph zweizusammenhängend wird. Wir sollten jedoch beim Einfügen der Kanten darauf achten, daß der resultierende Graph danach planar bleibt, falls der Eingabegraph planar war.

Das Problem, die minimale Anzahl von Kanten zu einem gegebenen planaren Graphen hinzuzufügen, so daß er planar bleibt und zweizusammenhängend wird, heißt planare Augmentierung. Auch für dieses Problem wird es mit hoher Wahrscheinlichkeit keinen effizienten Algorithmus geben, d.h. im komplexitätstheoretischen Sinne ist es NP-schwer. Bisher ist noch keines der in der Literatur angegebenen Probleminstanzen beweisbar zur Optimalität gelöst worden. Der Erfolg der Schnittebenentechniken für das Problem, einen größten planaren Untergraphen zu bestimmen, gibt jedoch Anlaß zur Hoffnung, daß es auch für dieses Problem gelingen könnte, Probleminstanzen moderater Größenordnungen bereits in naher Zukunft lösen zu können.

Statt nun zuerst einen gegebenen nicht-planaren Graphen zu planarisieren und diesen dann zu augmentieren, ist es von Vorteil, beide Schritte gleichzeitig durchzuführen. Gegeben ein Graph, wir suchen nach einem planaren zweizusammenhängenden Graphen, der dem gegebenen so ähnlich wie möglich ist, d.h. es darf eine minimale Anzahl von Kanten entweder entfernt oder neue Kanten hinzugefügt werden. Dieses neue Problem, das wir (Michael Jünger und Petra Mutzel) auf der "Graph Drawing'94", der jährlichen internationalen Konferenz über das Automatisierte Zeichnen von Graphen, vorschlugen, stieß dort auf reges Interesse.

Diese drei Probleme, das planare Augmentierungsproblem, das zweizusammenhängende, größte planare Untergraphenproblem und das neu eingeführte Ähnlichkeitsproblem, können ein und demselben Polyeder zugeordnet werden. Strukturelle Untersuchungen dieses Polyeders führen also automatisch zu einer Verbesserung der Lösbarkeit aller drei Probleme. Erste theoretische sowie praktische Ergebnisse liegen bereits vor.

Nach erfolgter planarer Augmentierung kann der neue, planare und zweizusammenhängende Graph gezeichnet werden, wobei die hinzugefügten Kanten in der Zeichnung unterdrückt werden.

Kreuzungsminimierung. Die zunächst entfernten Kanten sollen nun wieder in die Zeichnung bzw. in die topologische Einbettung eingefügt werden, so daß die Anzahl der sich kreuzenden Linien minimiert wird. Wir bezeichnen dieses Problem als "eingeschränktes" Kreuzungsminimierungsproblem, da die Einbettung eines Teils der Kanten bereits festgelegt ist. Das uneingeschränkte Kreuzungsminimierungsproblem ist eines derjenigen graphentheoretischen Probleme, über die sehr viel publiziert wird, deren Ergebnisse jedoch noch sehr gering sind. Bisher ist es nur für sehr spezielle Graphklassen gelungen, die minimale Anzahl der Kreuzungen von Kanten zu bestimmen. Selbst für einfache Graphklassen, wie zum Beispiel für vollständige Graphen ist nur eine Vermutung bekannt. Der Beweis dieser Vermutung steht schon seit über 30 Jahren aus. Wir hoffen, mit Schnittebenenverfahren zumindest für spezielle Graphenklassen auch hier etwas beitragen zu können, befinden uns jedoch gerade noch in den Anfängen unserer Forschung. Für das kreuzungsminimale Wiedereinfügen der zunächst entfernten Kanten müssen wir uns bisher mit Näherungsverfahren begnügen.

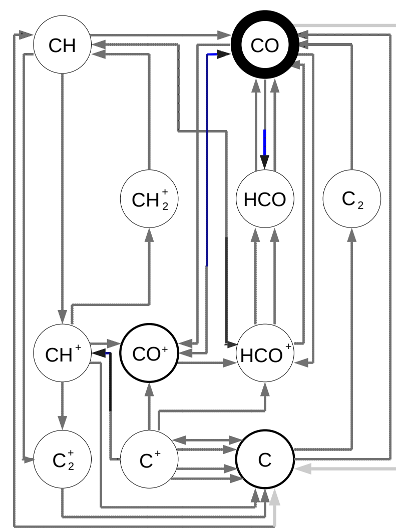

Abb. 6: Das mit der beschriebenen Methode (noch halbautomatisch) gezeichnete chemische Reaktionsdiagramm von Abbildung 1.

Die einzelnen Schritte dieser Methode sind bisher nur zum Teil implementiert. Abbildung 6 zeigt, wie das in Abbildung 1 gezeigte chemische Reaktionsdiagramm nach den beschriebenen Schritten aussehen könnte (*Mutzel*).