

Peter Gärdenfors: **Knowledge in Flux.**
Cambridge, Mass., London: Bradford Books, MIT
Press 1988, 262 Seiten, 36.95 \$.

Mit diesem Buch legt Peter Gärdenfors, Professor für Cognitive Science in Lund (Schweden), sein Modell der Dynamik von Überzeugungen ("beliefs") oder Theorien vor. Obwohl ein Großteil des Materials bereits in 15 Aufsätzen während der letzten zehn Jahre veröffentlicht worden ist, wirkt das Buch frisch, aktuell und wie aus einem Guß.

Der Aufbau des Buchs ist einfach und klar. Im ersten Teil beschäftigen sich fünf Kapitel mit den Modellen epistemischer Zustände. Die ersten beiden Kapitel gehen die Elemente epistemologischer Theorien und eine Reihe möglicher Repräsentationen von Wissen, oder besser: von Überzeugungen durch. Kapitel 3 und 4 diskutieren, wie sich Überzeugungen ändern, wenn man sie in Form von *konsistenten* und *deduktiv geschlossenen* Satzmengen (Theorien im Sinne des Logikers oder "belief sets") darstellt. Gärdenfors unterscheidet drei Grundformen solcher Änderungen:

- Expansionen K^+_A , bei denen zur alten Theorie K ein Satz A als Neuinformation oder Hypothese konsistent hinzugefügt werden kann;
- (minimale) Revisionen K^*_A , wo der aufzunehmende Satz A der ursprünglichen "Wissensbasis" K widerspricht; und
- Kontraktionen K^-_A , wo ein Satz A aus der ursprünglichen "Wissensbasis" K zu tilgen ist.

Während Expansionen kanonisch durchzuführen sind, bereitet bei Revisionen die Konsistenzbedingung und bei Kontraktionen die Abgeschlossenheitsbedingung Schwierigkeiten. Gärdenfors stellt zunächst Rationalitätspostulate für Revisionen und Kontraktionen auf und macht deutlich, daß sich die beiden Typen von Änderungen wechselseitig definieren lassen ($K^*_A = [K^-_A]^+_A$ bzw. $K^-_A = K \cap K^*_{-A}$). Deshalb genügt es im Anschluß zu zeigen, auf welche Weise man Kontraktionen konstruieren kann. In einigen Repräsentationstheoremen weist Gärdenfors nach, daß die Rationalitätspostulate durch diese Konstruktionen erfüllt und sozusagen ausgeschöpft werden. In Kapitel 5 überträgt Gärdenfors Terminologie und Resultate für Belief Sets auf probabilistische, Bayesianische Modelle der Wissensrepräsentation.

Im zweiten Teil des Buchs präsentiert Gärdenfors vier Anwendungen seiner Theorie der Überzeugungsänderungen. Kapitel 6 zeigt, wie man verschiedene Systeme der Aussagenlogik erzeugt, indem man Propositionen als Funktionen auffaßt, die epistemische Zustände als Argumente und Werte haben. Die Logik von Konditionalsätzen wird in Kapitel 7 über den sog. Ramsey-Test analysiert, nach welchem "Wenn A , dann B " genau dann in einem Belief set K ist, wenn B Element von K^*_A ist. Hier wird aber auch erstmals ein fundamentales Problem des Ramsey-Tests aufgezeigt. Unter Heranziehung probabilistischer Modelle werden in den letzten beiden Kapiteln

des Buchs schließlich die wissenschaftstheoretischen Fragestellungen der Erklärung und der Kausalität behandelt, wobei als entscheidende Idee in beiden Fällen ein Kriterium der positiven Relevanz bzgl. einer Kontraktion der aktuellen Wahrscheinlichkeitsfunktion dient.

Das Buch entstammt ganz unverkennbar der logisch-philosophischen Tradition. Es beinhaltet eine Fülle neuer Ideen, die auf stets stringente, formal präzierte Weise dargeboten werden. (Es gibt einen 30-seitigen Anhang mit Beweisen der wichtigsten Resultate). Gärdenfors volle Originalität zeigt sich vor allem im dritten, vierten und siebten Kapitel, die meines Erachtens das Herzstück des ganzen Buchs ausmachen. Die "Gärdenfors-Postulate" für Revisionen und Kontraktionen haben ebenso wie das "Gärdenforsche Unmöglichkeitstheorem" für den Ramsey-Test mittlerweile auch über die engeren Fachkreise hinaus Bekanntheit erlangt. In der Natur der Sache liegt, daß es gerade hier noch interessante Fragen gibt, die offen bleiben: Wie hängen die zwei *prima facie* völlig verschiedenen Konstruktionen von Kontraktionen aus den Abschnitten 4.4 und 4.8 zusammen? Wie konstruiert man iterierte Revisionen? Worin liegt der "tiefere Grund" für das Gärdenforsche Unmöglichkeitstheorem, und welche Konsequenzen sollen wir daraus ziehen?

Wenngleich der Autor nur beiläufig auf KI-Literatur (Doyle, Johnson-Laird, Fagin et al.) Bezug nimmt, so sind in dem Buch doch durchweg Dinge zu finden, die aus der KI-Perspektive höchst relevant erscheinen. So weiß man etwa nicht erst seit den Arbeiten von M. Ginsberg oder Delgrande, daß Konditionalsätze in engen Zusammenhang mit nichtmonotonen Schlußweisen und mit dem Frame-Problem zu bringen sind. Anzumerken ist allerdings, daß der im Buch verwendete Begriff der Nichtmonotonie vom Standard abweicht: Während man damit normalerweise meint, daß für *Axiomenmengen* K und H mit $K \subseteq H$ i.a. nicht $\text{Thm}(K) \subseteq \text{Thm}(H)$ gilt, heißt Nichtmonotonie bei Gärdenfors, daß für Theorien K und H aus $K \subseteq H$ i.a. nicht $K^*_A \subseteq H^*_A$ folgt. Gärdenfors diskutiert keine nichtmonotonen Konsequenzrelationen.

Für den KI-ler am meisten zu holen ist aus dem Buch zum Thema "Updating von Wissensbasen". Gärdenfors arbeitet zwar stets mit der Idealisierung konsistenter und deduktiv abgeschlossener Satzmengen, was natürlich eine unmittelbare praktische Verwertbarkeit (Implementierung!) seines Modells verhindert. (Ansätze, im Modell mit einer Unterscheidung zwischen explizitem und implizitem Wissen ohne diese Idealisierung auszukommen, kann man sowohl im Bereich der philosophischen Logik als auch der KI ausmachen.) Insofern als die hier vorgelegte Theorie uns sagt, was passieren sollte, wenn unsere wissensbasierten Systeme eine perfekte Logikkomponente hätten, ist sie jedoch sehr attraktiv. Als eigenständige "Kohärenztheorie" der Wissensrevision bietet sie eine echte Alternative zu den "Begründungstheorien" (diese Unterscheidung stammt von

Harman) der gängigen Reason-Maintenance-Systeme.

Dieses Buch kann jedem, der sich mit den angesprochenen Themen auseinandersetzt, nur wärmstens empfohlen werden.

Hans Rott, München
