



Technische Universität München

Fakultät für Informatik

Forschungs- und Lehrereinheit Informatik VI

Überblick kognitiver Architekturen

Seminar Kognitive Robotik (SS12)

Iris Rothkirch

Sebastian Hofstetter

Betreuer: Dr. Florian Röhrbein

Leitung: Prof. Alois Knoll

Abgabetermin: 21. Juli 2012

Inhaltsverzeichnis

1	Einführung	3
2	Kognition und Intelligenz.....	3
2.1	Begriffsklärung	3
2.2	Informationsverarbeitungsgesetz	3
3	Kognitive Architekturen.....	5
3.1	Begriffsklärung	5
3.2	Herangehensweisen und Klassifizierungen	5
3.3	Strukturen.....	5
3.3.1	Informationsarten	5
3.3.2	Informationsspeicher	6
3.3.3	Informationsverarbeitungsprozess	7
4	Funktionale Analyse	9
4.1	Kognitives Spektrum.....	9
4.2	Verhaltensspektrum	9
5	Überblick über Architekturen Modelle	9
5.1	SOAR.....	9
5.1.1	Die Gesamtarchitektur	9
5.1.2	Produktionsregeln	10
5.1.3	Operatoren.....	10
5.1.4	Zustand	10
5.1.5	Entscheidungszyklus.....	11
5.1.6	Sackgassen.....	11
5.1.7	Lernmechanismen	12
5.2	ICARUS.....	12
6	Gefahren und Risiken der Künstlichen Intelligenz.....	13
7	Literaturverzeichnis.....	14
8	Abbildungsverzeichnis	15

1 Einführung

Die Modellierung kognitiver Prozesse ist zentraler Bestandteil kognitionswissenschaftlicher Forschung. Grundlegend dabei ist die Frage, wie Organismen einfache und komplexe Prozesse der Informationsverarbeitung bei der Wahrnehmung, beim Denken und bei Handlungen in Interaktion mit der Umwelt organisieren, ausführen und kontrollieren. Die interdisziplinäre Kognitionswissenschaft liefert hierzu verschiedenen Erklärungsansätze, die sich in ihren begrifflichen Konzepten, ihrem Modellierungsanspruch und in der Architektur der Modelle zum Teil gravierend unterscheiden. Die Kognitionswissenschaft hat sich aus einem Spannungsfeld zwischen Kognitionspsychologie, Computerwissenschaften (Künstlicher Intelligenz) und den Neurowissenschaften entwickelt. Gemeinsam ist ihnen das Ziel herauszufinden „ wie der menschliche Geist funktioniert“. Allerdings favorisiert die natur- und neurowissenschaftliche Seite dabei eine reduktionistische Sichtweise, mit dem Ziel kognitive Phänomene auf der elementaren Ebene zu untersuchen. Wohingegen die Kognitionspsychologie an der Darstellung auf der repräsentationalen Ebene festhält, da diese für unverzichtbar angesehen wird, um Phänomene wie Wahrnehmung, Problemlösen, Sprache zu erklären. Somit ist ein interdisziplinärer Austausch zwischen diesen Teilgebieten notwendig.⁸

2 Kognition und Intelligenz

2.1 Begriffsklärung

Kognition

Im Grundsatz der theoretischen Philosophie Descartes- Ich denke also bin ich- wird ein wesentlicher Grundstein für die Selbstständigkeit im Denken, der Grundannahme eines kognitiven Prozesses gelegt. Aus dem Lateinischen „cognoscere“ stammend charakterisiert der Begriff Kognition allgemeine Prozesse des Wahrnehmens, der Erkenntnis, des Vorstellens, des Wissens, des Denkens, der Kommunikation und der Handlungsplanung. Kurz gesagt handelt es sich hierbei um die Gesamtheit aller Vorgänge, welche der Aufnahme, Verarbeitung und Speicherung von Informationen dienen. Von einem kognitiven System zu sprechen setzt voraus dass ein System relevante Aspekte der kognitiven Funktionen und/ oder Strukturen der vorgegeben Definition aufweist.^{2,8}

Intelligenz

In der Psychologie hat man sich bis heute nicht auf eine allgemein akzeptierte Definition geeinigt. Allerdings sind wichtige Kriterien der Intelligenz zum einen das zielgerichtete Verhalten und zum anderen die Rationalität.⁸

Zusammenfassend lässt sich sagen, dass Kognition in der (intelligenten) Verarbeitung von Informationen zur Steuerung des Verhaltens besteht. Und Intelligenz die zweckmäßige und erfolgreiche Kognition in Bezug auf Wissen darstellt.

2.2 Informationsverarbeitungsgesetz

Der Verarbeitung kommt dabei im Kontext der Kognitiven Architekturen eine besondere Bedeutung zu. Der Grundgedanke des Informationsverarbeitungsgesetzes besteht in der Annahme den Menschen als informationsverarbeitendes System zu betrachten.

Der bekannte Ablauf „Input -> System -> Output“ beim Computer spiegelt sich beim Menschen in der Abfolge „Wahrnehmung -> Bewusstsein -> Motorik“ wieder.

Belege dafür dass das menschliche System dem Computer gleich gesetzt werden kann sind folgende:

1. Information lässt sich durch Symbole repräsentieren.
2. Informationsverarbeitung lässt sich durch formale Operationen der Symbole beschreiben.
3. Formale Operationen sind zunächst abstrakt und unabhängig von ihrer materiellen Grundlage.¹¹

Dies würde den Einsatz von Computerprogrammen zu Simulation kognitiver Prozesse und die Möglichkeit von künstlicher Intelligenz auf Computern rechtfertigen.⁸

Es gab Kognitionswissenschaftler die die Computeranalogie wörtlich nahmen und versuchten zu zeigen, dass die menschliche Kognition nicht nur einem Computer vergleichbar ist sondern tatsächlich ein Computer ist, man spricht hier auch von der starken Künstlichen Intelligenz.

1. Physical Symbol System Hypothesis

Allen Newell und Herbert A.Simon entwickelten 1975 die PSS Hypothese. Ihre Kernaussage war folgende: *„Ein physikalisches symbolverarbeitendes System besitzt die notwendigen und hinreichenden Voraussetzungen für allgemeines intelligentes Handeln.“*

Sie vertraten die Auffassung, dass ein materiell realisiertes Symbolverarbeitungssystem alle notwendigen und hinreichenden Voraussetzungen für Kognition besitzt.

Die These impliziert somit zweierlei:

1. Die menschliche Intelligenz beruht auf der Manipulation von Symbolen („notwendige" Seite).
2. Jedes System, das die Fähigkeit zur Symbolmanipulation besitzt, ist in der Lage, Intelligenz zu entwickeln („hinreichende" Seite).¹²

Das heißt, PSS (Computer) können intelligentes Verhalten zeigen und intelligentes Verhalten (von Menschen) lässt sich durch PSS beschreiben.²

2. Unified Theories of Cognition

Angestoßen von Allen Newell entwickelte sich hinter diesem Begriff die Idee, die komplette Bandbreite menschlicher kognitiver Fähigkeiten in einer einzigen, großen Theorie zu vereinigen. Somit ist sie weniger eine einzelne Theorie als eine ganze Familie von Theorien mit der Ansicht dass der menschlichen Geist vom Prinzip her auf die gleicher Art funktioniert wie ein Computer und folglich auf der funktionalen Ebene ein Computer ist. Als ein wichtiger Bestandteil menschlicher Intelligenz wird hier eine Wissensbasis gesehen, d.h. man sieht den Menschen als „Wissenssystem“. Als eines der wichtigsten Konzepte der Wissensrepräsentation wird dabei das Symbol gesehen, da es die Möglichkeit bietet verschieden Objekte des Denkens miteinander in Verbindung zu bringen. Newell ist also von der Möglichkeit überzeugt ein einziges System von Mechanismen in einer kognitionswissenschaftlich fundierten Theorie hervorzubringen, das die gesamte Bandbreite von Kognition abdeckt.^{4,2}

Als eigenen Kandidaten für eine Unified Theory nennt Newell das Symbolsystem SOAR, im späteren wird noch genauer auf diese kognitive Architektur eingegangen.⁸

3 Kognitive Architekturen

3.1 Begriffsklärung

Im Kontext der Informationsverarbeitung versteht man unter der Architektur die Menge der Mechanismen, die zum Erlangen kognitiver Fähigkeiten befähigen. Diese sind im Rahmen einer Spezifikation erfasst und geben den daraus abgeleiteten kognitiven Systemen eine feste gemeinsame Struktur.¹⁵

3.2 Herangehensweisen und Klassifizierungen

Kognitivistische und emergente Ansätze

Zwei grundlegend philosophisch verschiedene Herangehensweisen werden durch *kognitivistische* (*cognitivist*) und *emergente* (*emergent*) Architekturen verkörpert. Während Erstere dem Prinzip der Informationsverarbeitung folgt und worauf auch aktuell der Fokus liegt, gehen emergente Architekturen davon aus, dass Intelligenz spontan aus dem Zusammenschluss der (primitiven) Komponenten *auftauchen* (*emergieren*) kann.¹⁵

Interne Modellierungen

Kognitive Architekturen lassen sich auch anhand der internen Vernetzung der Teilkomponenten in integrierende, geschichtete, modulare und konnektionistische Architekturen aufteilen.¹⁵

3.3 Strukturen

Gemäß dem Informationsverarbeitungsgesetz besteht die Funktion kognitiver Prozesse in der Transformation von Informationen. Eine Architektur ist somit charakterisiert durch die Informationsarten, Möglichkeiten der Speicherung, sowie die verschiedenen Verarbeitungsprozesse.

3.3.1 Informationsarten

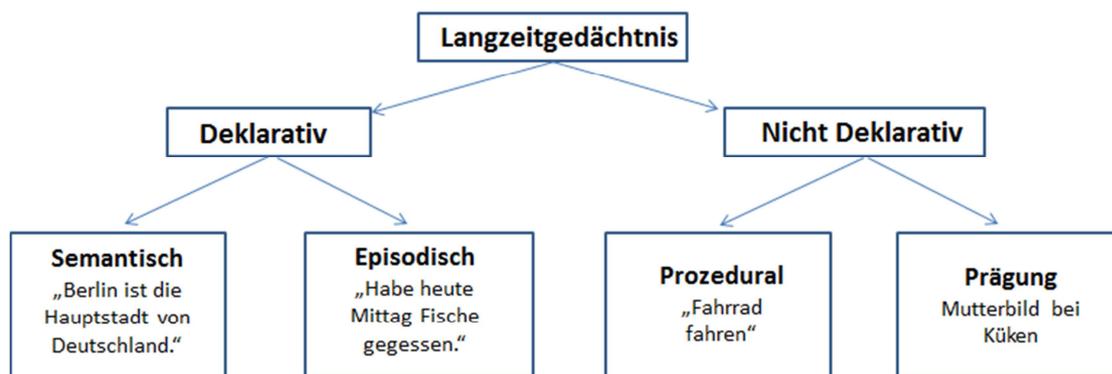


ABBILDUNG 1: ÜBERBLICK ÜBER WISSENSARTEN

Deklaratives Wissen: Die deklarative Information stellt die Repräsentationen der Welt in Form von Fakten dar. Sie ist episodisch, wenn sie von konkreten Einzelereignissen handelt und semantisch, wenn sie allgemeine und kontextunabhängige Tatsachen beschreibt. In kognitiven Architekturen wird deklaratives Wissen durch symbolische Ausdrücke beschrieben. In den meisten Architekturen werden Objekte und Fakten über Objekte durch Eigenschaftslisten repräsentiert.^{2, 10}

Prozedurales Wissen: Prozedurale Information beinhaltet kognitive und motorische Fähigkeiten. Während deklaratives Wissen abstrakt und unabhängig von der jeweiligen Architektur ist, wird prozedurales Wissen direkt von der Architektur ausgeführt. Deklaratives Wissen kann nur verwendet werden, indem es prozedural ausgeführt wird. Im prozeduralen Wissen werden Ableitungsregeln definiert, mit deren Hilfe aus bereits bestehenden Fakten neue Fakten gewonnen werden.²

3.3.2 Informationsspeicher

3.3.2.1 Speichermodule

Jeder kennt die Situation, dass man im Telefonbuch eine Nummer anschaut, zum Telefon greift, wählt, aber nach 2 Minuten ist die Nummer dann auch schon wieder vergessen. So betrachtet stellt das Bewusstsein eine Art Gedächtnis dar, in dem Information direkt verfügbar ist, aber offensichtlich auch schnell verloren gehen kann. In diesem Sinne wurde bereits im 19. Jahrhundert die Theorie eines Primär- und Sekundärgedächtnisses entwickelt. Diese Unterscheidung zwischen Kurz- und Langzeitgedächtnis wurde zum Kern des so genannten Multi-Speicher Modells von Atkinson und Shiffrin (1968). Nach diesem Modell wird einlaufende Information kurzfristig im sensorischen Register gehalten und die beachtete Information findet Eingang in das Kurzzeitgedächtnis, dessen Kapazität allerdings beschränkt ist. Hier zerfällt die Information, außer sie wird in ausreichendem Maße memoriert, d.h. wiederholt, sodass sie Zugang ins Langzeitgedächtnis findet. Allerdings kam schnell Kritik gegenüber diesem Modell auf, da festgestellt wurde, dass beim Verstehen eines Satzes Informationen aus dem LZG verwendet werden. Deshalb wurde ein neues Modell von Baddeley 1974 aufgestellt, bei dem es anstelle von einem KZG ein Arbeitsgedächtnis gibt, welches die augenblicklichen Inhalte des Denkens enthält („Desktop“). Ein dritter Ansatz wurde von Cowan und Oberauer aufgestellt, die behaupteten, dass es eigentlich überhaupt kein separates KZG bzw. AG gibt, sondern vielmehr nur ein LZG welches ein riesiges Netzwerk an verbundenen Informationen ist. Gewisse Teile können davon aktiviert sein, die dann quasi das Arbeitsgedächtnis darstellen.^{10,11}

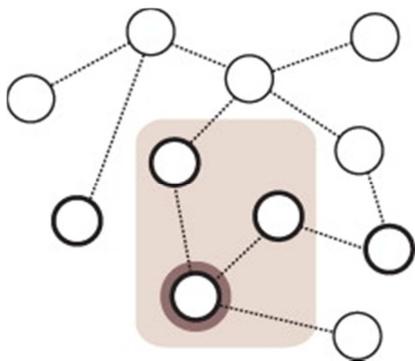


ABBILDUNG 2: COWAN-OBERAUER MODELL¹⁴

In Bezug auf kognitive Architekturen unterscheidet man nun Speichermodule für langfristige und kurzfristige Speicherung. Gerade bei großen Datenmengen ist es sinnvoll, einen längerfristigen Speicher von einem Arbeitsspeicher zu trennen, dessen Inhalte die momentan kognitiven Prozesse direkt nutzen können, jedoch entfernt werden, sobald sie nicht mehr gebraucht werden. Folglich unterscheidet man zwischen einer langfristigen Speicherung für große Datenmengen und einer kurzfristigen Speicherung für Informationen, die schnell verfügbar sein müssen.²

3.3.2 Speicherorganisation und Zugriff

Neben der Aufteilung in Speichermodule ist auch die Organisation der Informationen innerhalb eines Moduls wichtig, sodass man einen möglichst effizienten Zugriff erreicht und ein schneller Abruf möglich ist. SOAR und ACT-R lösen dieses Problem beispielsweise mit Assoziationsmechanismen. Assoziationen verbinden Speicherinhalte miteinander, die in einem Bezug stehen und oft zusammen benötigt werden. Somit werden die Informationen aus dem LZS in den Arbeitsspeicher übertragen, wenn diese in der gegebenen Situation nützlich sein könnten. Der Vorteil des assoziativen Speicherzugriffs besteht in einem bedeutungsadressierten Speicherzugriff, bei dem keine vollständige und aufwändige Suche nach Informationen durchgeführt werden muss. Allerdings stellt die fehlende Möglichkeit einer direkten Suche einen Nachteil dar, also das Problem wenn Information gebraucht wird, die in keiner assoziativen Verbindung steht.

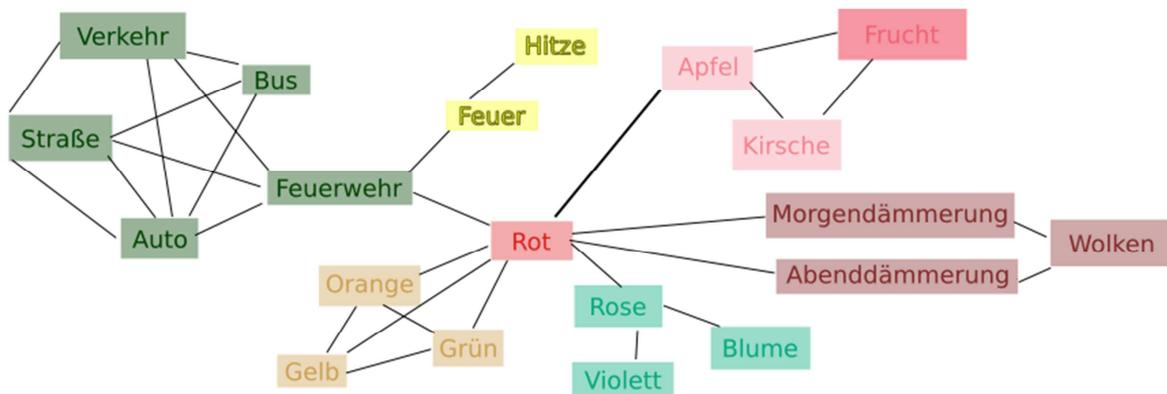


ABBILDUNG 3: AKTIVIERUNGS-AUSBREITUNG MIT ASSOZIATIONSMECHANISMEN¹¹

3.3.3 Informationsverarbeitungsprozess

Abbildung 4 stellt schematisch die grundlegenden Informationsverarbeitungsprozesse und den Informationsfluss zwischen ihnen dar. Information muss aufgenommen, gespeichert, aus dem Speicher abgerufen, transformiert und schließlich für Handlungen genutzt werden können. Daraus ergibt sich der folgende Kreislauf aus Wahrnehmung, Denken und Handeln²

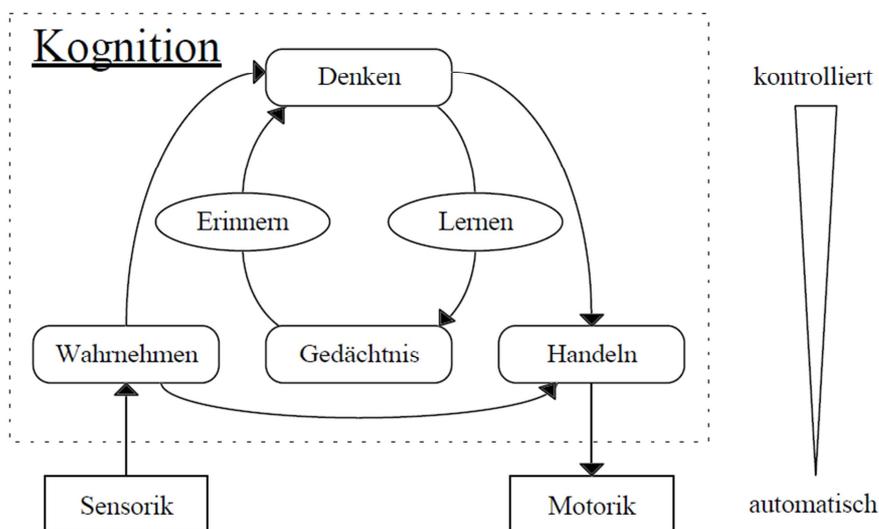


ABBILDUNG 4: INFORMATIONSVERARBEITUNGSPROZESS²

Die zentrale Frage ist nun wie ein Agent zu seinem Wissen gelangt. Viele der Fähigkeiten, die im Folgenden diskutiert werden, geben dem Agenten Zugang zu solch einem Wissen.

1. Wahrnehmen und Handeln

Wahrnehmung und Handeln stellen die Schnittstelle zwischen dem Agenten und seiner Umwelt dar. Kognition geschieht nicht in Isolation, sondern ein intelligenter Agent existiert immer in einem Kontext einer externen Umwelt, den er spüren, wahrnehmen und interpretieren muss. Dies kann man mit einem Menschen vergleichen, der verschiedene Sinnesmodalitäten wie Sehen, Hören und Fühlen hat. Aus der Wahrnehmung gewinnt er sein Wissen über die Welt, durch das Handeln kann er diese verändern.^{1,2,4}

2. Ausführung und Aktion

Kognition findet statt um Aktivität in der Umwelt zu unterstützen und auszuführen. Deshalb muss eine Architektur motorische Fähigkeiten besitzen, um diese Aktivitäten zu ermöglichen. Zum Beispiel sollte ein mobiler Roboter sich von einem Platz zum anderen bewegen können.^{1,2,4}

3. Erinnern und Lernen

Unter Erinnern versteht man die Fähigkeit zu enkodieren und das Ergebnis eines kognitiven Prozesses zu speichern, um später dazu Zugang zu haben. Ein Agent kann sich nicht wirklich an seinen eigenen physischen Aktionen erinnern, vielmehr kann er die kognitiven Strukturen wieder aufrufen, die diese Handlung beschrieben haben.

Lernen beinhaltet hauptsächlich die Generalisierung von spezifischen Wissen und Handlungen.^{1,2,4}

4. Problemlösen und Planen

Unter Problemlösen versteht man die kognitiven Tätigkeiten, mit denen Handlungen zum Verändern der Welt auf ein bestimmtes Ziel oder einen Zweck hin erzeugt werden. Da das Handeln kognitiver Architekturen immer zweckgerichtet ist, lässt sich alle kognitive Aktivität, sofern sie sich mit Handlungen befasst, als das Lösen von Problemen auffassen.

Während Problemlösen auch in der Welt stattfinden kann, findet die Planung normalerweise im Kopf des Agenten statt. Eine kognitive Architektur muss fähig sein ihren Plan als eine geordnete Folge von Aktionen zu präsentieren. Man unterscheidet dabei drei verschiedene Ebenen der Planung. Bei der reaktiven Planung wird durch eine wahrgenommen Situation direkt eine angemessene Reaktion ausgelöst; die situative Planung berücksichtigt den situativen Kontext; die deliberative Planung geschieht völlig unabhängig von äußerlichen Veränderungen. Diese drei Arten schließen einander nicht aus, sondern sollten am besten in einer Architektur kombiniert werden.^{1,2,4}

5. Logisches Denken und Entscheiden

Problemlösen steht in engem Zusammenhang zu logischem Denken, eine weiter zentrale kognitive Fähigkeit, welche den Agenten seinen Wissensstand vergrößern lässt. Während Planen hauptsächlich das Erreichen von Zielen in der Welt betreffen, zieht das logische Denken geistige Schlussfolgerungen aus Wissen oder Annahmen über die der Agent bereits verfügt. Man unterscheidet zwischen deduktiven Schlussfolgerungen, bei der nach logischen Regeln sichere Konsequenzen aus dem vorhandenen Wissen gezogen werden und induktive Schlussfolgerungen, die unsicher sind, da sie von einzelnen Fällen auf die Allgemeinheit schließen.

Entscheiden bedeutet die zielgerichtete Auswahl von Alternativen. Solche Entscheidungen sind oft mit dem Wiedererkennung einer Situation oder einem Muster verbunden.^{1,2,4}

6. Interaktion und Kommunikation

Manchmal ist der effektivste Weg für einen Agenten um Wissen zu erlangen, indem er mit einem anderen Agenten kommuniziert. Agenten existieren in einer Umwelt mit anderen Agenten insofern gibt es viele Gelegenheiten in welchen sie wissen transferieren müssen.^{1,2,4}

4 Funktionale Analyse¹⁵

4.1 Kognitives Spektrum

1. Reaktivität

Der Agent sollte jederzeit in der Lage sein und möglichst in Echtzeit zu reagieren.

2. Situativität

Durch hohes Abstraktionsvermögen werden in unplanmäßige Situationen adäquate Lösungsmöglichkeiten erreicht.

3. Deliberativität

Das willkürliche Verhalten gemäß einer internen Zielsetzung wird häufig noch einmal separat betrachtet.

4. Adaptivität

Sie kann zunächst als Kombination aus Reaktivität und Situativität betrachtet werden.

5. Sozialität

Durch den Informationsaustausch mit anderen Agenten lassen sich Vorteile erzielen (*Schwarmverhalten*).

4.2 Verhaltensspektrum

1. Rationalität

Rationalität lässt sich anhand des deliberativen Verhaltens bewerten und lässt sich in Effizienz, Kohärenz von Wissen und Handlung und Kohärenz des Gesamtverhaltens aufteilen.

2. Robustheit

Eine Architektur ist umso robuster, je besser sie mit Widrigkeiten wie z.B. unvollständigem Wissen umgehen kann.

3. Flexibilität und Vielseitigkeit

Allgemein ist es anzustreben eine maximal vielseitige Architektur zu entwerfen, die sich flexibel auf einzelne Problemfelder spezialisieren kann.

5 Überblick über Architekturen Modelle

5.1 SOAR

5.1.1 Die Gesamtarchitektur

Die Architektur wurde 1983 von A. Newell, J. Laird und P. Rosenbloom entwickelt. Seitdem hat sie viele Entwicklungen durchgemacht bis jetzt zu aktuellen Version 9.3.2, welche auf der Homepage (<http://sitemaker.umich.edu/soar/home>) frei zum Downloaden verfügbar ist. Der Name SOAR ist eine Abkürzung für State, Operator And Result. Im Groben stellt der Name somit die Hauptteile von SOAR dar. Die Problemlösung in SOAR wird als Suche in einem Problemraum angesehen, in dem man Operatoren an Zustände anwendet um Ergebnisse zu erzielen. Die Problemraumsuche findet in einem zentralen Arbeitsspeicher (working memory) statt, der das temporäre Wissen beinhaltet. Damit Wissen verwendet werden kann, muss es zuvor aus dem Langzeitspeicher in den Arbeitsspeicher abgerufen werden. Das Wissen im Langzeitspeicher (production memory) ist assoziativ durch Produktionen gespeichert, die Wissensseinheiten in den Arbeitsspeicher schreiben (execute), wann immer ihre Vorbedingung erfüllt ist (match). Weiterhin sind mit dem Arbeitsspeicher Systeme Das Wissen im Langzeitspeicher entsteht durch permanentes Lernen aus Erfahrung (chunking).

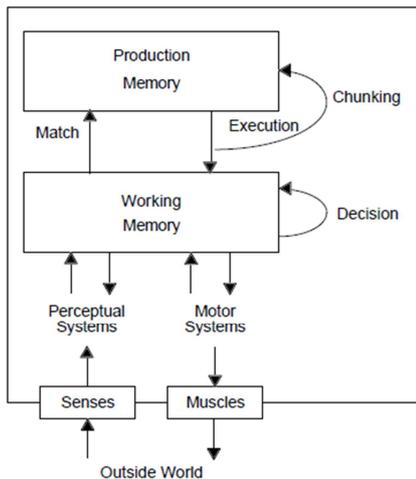


ABBILDUNG 5: ARCHITEKTUR VON SOAR²

5.1.2 Produktionsregeln

Jede Regel beginnt mit „sp“: „SOAR Production“. Danach folgt der Körper der Regel. Dieser enthält den Regelnamen, gefolgt von mindestens einer Bedingung, dem Pfeilsymbol und mindestens einer Aktion. Eine Bedingung testet das Vorhandensein von Daten im Arbeitsspeicher. Wenn alle Bedingungen zutreffen feuert die Regel und die Aktion wird ausgelöst.⁶

Quelltext	Kommentar
sp { propose*hello-world	world # Diese Regel schlägt vor, den Operator "hello-world" anzuwenden
(state <s> ^type state)	# Bedingung: WENN im Arbeitsspeicher ein Zustand <s> existiert, dann "feuert" die Regel (Aktion 1 und 2 werden ausgeführt)
→	
(<s> ^operator <o> +)	# Aktion1: schlägt vor, auf den momentanen Zustand einen Operator <o> anzuwenden
(<o> ^name hello-world)}	# Aktion2: <o> bekommt den Namen "hello-world"

5.1.3 Operatoren

Wie oben bereits gesehen können Regeln in SOAR spezielle Aktionen erzeugen. Jede ausgeführte Aktion erzeugt Operatoren im Arbeitsspeicher. Diese gelten als Vorschläge, die durch einen Entscheidungsalgorithmus bestimmt werden. Eine Regel bietet einen Operator an, indem sie den Operator mit einer Präferenz versieht („+“).⁶

5.1.4 Zustand

Mit den Produktionsregeln werden Speicherobjekte verändert. Im Folgenden wird betrachtet wie Objekte in SOAR modelliert werden. Wissen wird durch Attribut- Wert Listen repräsentiert und in Form von Produktionen gespeichert. Im Arbeitsspeicher werden einzelne Wissensobjekte in einem Zustand zusammengefasst. Ein Zustand enthält alle Informationen über die momentane Situation. Ein Zustand besitzt folglich Werte und Attribute, wobei die Werte wieder Attribute besitzen können. In SOAR ist das ganze Wissen in einem Graphen organisiert.⁶



ABBILDUNG 6: REPRÄSENTATION EINES ZUSTANDES⁶

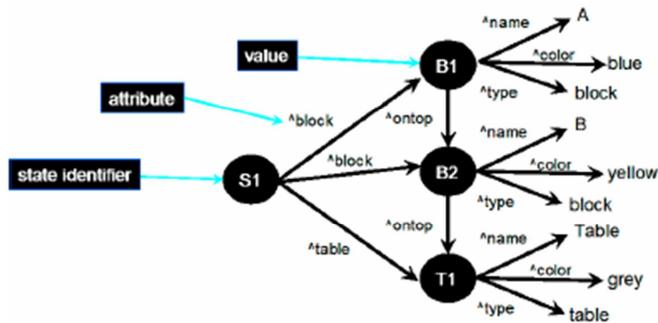


ABBILDUNG 7: REPRÄSENTATION EINES ZUSTANDES⁶

5.1.5 Entscheidungszyklus

Er besteht aus der Phasen:

1. Input (optional)
2. Wissensuche (Elaborationsphase):

Die Vorbedingungen der Produktionen, die den Langzeitspeicher bilden, werden im WM überprüft und deren Wissen, falls die Bedingung erfüllt ist, zum WM hinzugefügt.

3. Select Operator:

Es wird die eigentliche Entscheidung getroffen anhand des im WM enthaltenen Wissens.

1. Fall: Ein Operator hat eine höhere Präferenz als alle anderen, dann wird dieser angewendet.
2. Fall: Es kann nicht genau ein Operator ausgewählt werden (-> Sackgasse).

4. Apply Operator State:

Es werden die Aktionen ausgelöst, die als Bedingung den ausgewählten Operator haben.

5. Output (optional)⁶

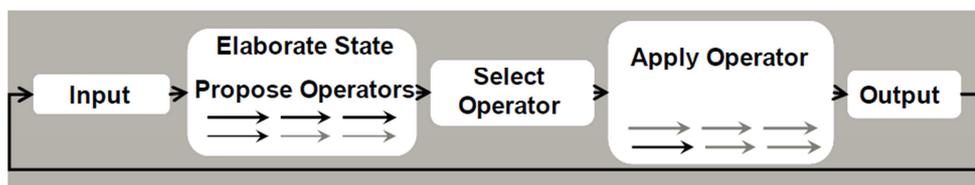


ABBILDUNG 8: ENTSCHEIDUNGSZYKLUS⁶

5.1.6 Sackgassen

Wie bereits erwähnt entstehen Sackgassen wenn die Entscheidung zwischen Operatoren nicht getroffen werden kann, d.h. sie besitzen die gleiche Priorität. Es wird dann automatisch ein neues Teilziel aufgestellt und zum Zielstapel hinzugefügt um diese aufzulösen. Diese Teilziele werden aus dem Zielstapel entfernt, sobald die Sackgasse aufgelöst wird, d.h. sobald eine Lösung für die Ursache des Problems gefunden ist.⁶

5.1.7 Lernmechanismen

Lernen durch Chunking geschieht permanent während des Problemlösevorgangs immer dann, wenn eine Sackgasse aufgelöst wird. Das Chunking erzeugt Regeln die als Bedingung die Operatoren beinhalten, zwischen denen SOAR keine Wahl treffen konnte, und alle Bedingungen, die während der Auflösung in der Sackgasse benutzt wurden. Anhand dieser Vorbedingungen kann in allen Situationen, in denen der gleiche Problemlösevorgang möglich ist, die Produktion angewendet werden. Als Inhalt der Produktion wird das Ergebnis der Auflösung der Sackgasse verwendet.⁶

Reinforcement Learning bezieht sich auf die Prioritäten zwischen Operatoren. Wenn man ein Ziel Erreicht, kann man dies als positives Feedback auffassen, wenn nicht, dann als negatives Feedback. In diesem Fall würde man die Priorität des letzten Operators verringern, damit in einer zukünftigen Entscheidung ein anderer Operator eher gewählt wird.^{1,2}

Episodic und semantic Learning akquirieren den jeweiligen Typ des prozeduralen Wissens.¹

5.2 ICARUS

ICARUS ist eine der neueren kognitiven Architekturen und wurde 2004 von Langley, Cummings und Shapiro entwickelt. Wissen wird hier entweder als Konzept (concept) oder Fähigkeit (skill) eingeordnet. Erstere repräsentieren Umgebungssituationen in Abhängigkeit von anderen Konzepten bzw. Wahrnehmungen (percepts). Fähigkeiten beschreiben eine Anleitung zum Lösen eines Problems (goal), indem diese wiederum in Teilprobleme zerlegt werden.¹⁵

Arbeitszyklus

Jeder Zyklus folgt dem *recognize-act* Prinzip: Zunächst werden per Mustererkennung gefundene Wahrnehmungen in einen Buffer geschrieben und mit bekannten atomaren Konzepten verglichen. Übereinstimmungen werden als Vermutungen (*beliefs*) in einem Kurzzeitspeicher abgelegt. Nachdem iterativ auch abstraktere Konzepte erkannt wurde, wird nun hierarchisch absteigend im Fähigkeitenbaum nach anwendbaren Fähigkeiten gesucht, die noch unerfüllte Teilprobleme lösen können. Auf diese Weise wird Schritt für Schritt das Hauptproblem gelöst.¹⁵

Organisation

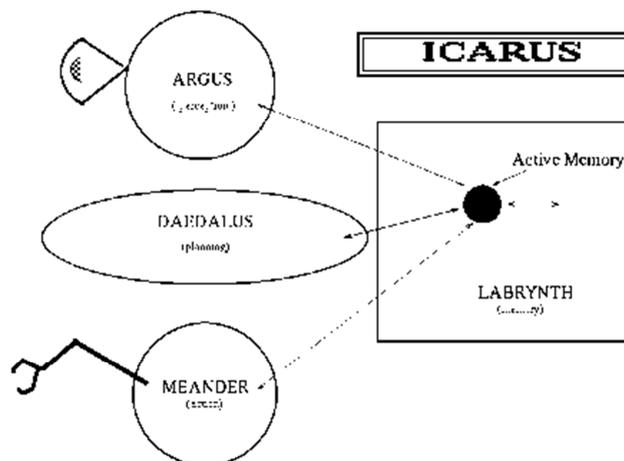


ABBILDUNG 9: HAUPTMODULE VON ICARUS

Der Argus (Wahrnehmung) generiert die Objekte aus der Umgebung. Dabei orientiert er sich am aktuellen Plan von Daedalus und setzt seine Ressourcen entsprechend ein. Außerdem kann ein spezielles Event an das Labyrinth gesendet werden, wenn die Umgebung außergewöhnliche Veränderung zeigt. Daedalus (Planung) führt die Planung für die Probleme durch, die im Labyrinth abgelegt wurden. Das Labyrinth (Speicher) unterscheidet sich von anderen Architekturen wesentlich dadurch, dass es auf eine Hierarchie probabilistischer Konzepte aufgebaut ist. Der Maender (Output) führt die Pläne von Daedalus aus.¹⁵

6 Gefahren und Risiken der Künstlichen Intelligenz

Bei aller Begeisterung für künstliche Systeme und der Idee das kognitive System Mensch zu entschlüsseln, darf man nicht die Gefahren und Risiken außer Acht lassen, die diese Forschungen mit sich bringen. Gerade ethische Aspekte werden in der Technik gerne in den Hintergrund gedrängt. Der Drang nach Wissen liegt nun mal in der Natur des Menschen, allerdings stellt sich hier die Frage, ob es diesmal nicht zu weit geht. In immer mehr Bereichen finden die Roboter und Automatisierungsprozesse Einzug, sei es in der Industrie oder im alltäglichen Leben. Und wo nichts als menschliche Perfektion erwartet wird haben die Menschen im Vergleich zu Robotern schlechte Karten, sie sind schneller, werden nicht müde und machen keine Fehler. Täglich gewöhnen uns die künstlichen Dienstleister ein Stück mehr an ihre Perfektion: Einkaufen im Internet, Fahrkartenverkauf und Flaschenrückgabe am Automaten. Dieser schleichende Automatisierungsprozess nimmt den Menschen ihre Individualität: die einzigen Kriterien sind Schnelligkeit und Zuverlässigkeit, man erwartet kein Lächeln, kein persönliches Wort mehr. Kann der Roboter also zum Feind des Menschen werden?

Ganz zu Beginn wurde das Informationsverarbeitungsgesetz angesprochen, dass besagt dass das menschliche System als Computer aufgefasst werden kann. Allerdings gibt es hier auch Kritikpunkte und zwar, dass damit nicht alle wichtigen Aspekte menschlicher Kognition angemessen bearbeitet sind. Gerade die Behandlung von Emotionen und Stimmungen, sowie das Verständnis und das Bewusstsein scheinen sich damit nicht gut vereinbaren zu lassen. Speicherkapazität und Verarbeitungsgeschwindigkeit sind nicht alles, was das menschliche Gehirn, was den Menschen, ausmacht. Wird ein Computer jemals Unbehagen, Freude, Trauer, Liebe, Hass verspüren? Wohl kaum. Wozu auch? Die Domäne des Computers ist nun mal die Logik, die des Menschen ist das Leben. Der Computer mag logisch kreativ sein, menschliche Kreativität aber ist der Gefühlswelt erschlossen.

Eines hat uns die Forschung zu mindestens im Bereich der Künstlichen Intelligenz gelehrt, wir können stolz auf uns, den Menschen, sein. Denn die Geschichte der Künstlichen Intelligenz hat uns immer wieder gezeigt wie unglaublich kompliziert der Mensch ist, wie schwer es ist, künstliche Intelligenz zu verstehen und zu erschaffen.^{7,9}

7 Literaturverzeichnis

- 1) Langley, P., Laird, J. E., & Rogers, S. (2008). Cognitive architectures: Research issues and challenges
Cognitive Systems Research, pp. 1-20.
- 2) Sessler, R. (1996). *Das Agentenmodell INTERAP im Vergleich mit kognitiven Architekturen*.
Professor J. Siekmann.
- 3) Vernon, D., Metta, G., & Sandini, G. (2007, April). A Survey of Artificial Cognitive Systems:
Implications for the Autonomous Development of Mental Capabilities in Computational Agents.
IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION(11/2), pp. 151-180.
- 4) Christian Michalk (2006/07): Eine Einführung in Kognitive Architekturen;
Dr. Michael Kipp, Dr. Alassane Ndiaye, Dr. Dominik Heckmann, Michael Feld
- 5) Kognitive Modellierung: Hauptseminar im WS 2002/2003
Joachim Funke, Psychologisches Institut, Uni Heidelberg
- 6) Peter Paul Kaczmarczyk: SOAR Eine Kognitive Architektur (2006/07)
Dozenten: Dr. Michael Kipp, Dr. Alassane Ndiaye, Dr. Dominik Heckmann
- 7) Maschine, Körper, Geist- eine Einführung in die Kognitionswissenschaft von Max Urchs, Vittorio
Klostermann GmbH 2002
- 8) Konnektionismus und Kognition von Markus Pospeschill, Verlag W. Kohlhammer , 2004
- 9) Wie denkt der Mensch? Informationstechnik im Gehirn von Peter R. Gerke, J.F. Bergmann Verlag
München, 1987
- 10) Lehrbuch der Kognitiven Psychologie von Johannes Engelkamp und Hubert D. Zimmer, Hogrefe
Verlag, 2006
- 11) Unterlagen Bachelorstudium Psychologie an der LMU "Allgemeine Psychologie 1"
- 12) http://www.lr-develop.de/wiki/index.php/Die_Physical_Symbol_System_Hypothesis_und_ihre_Geschichte
- 13) Künstliche Intelligenz von Roman Jerger, Hamburg
- 14) <http://www.sciencedirect.com/science/article/pii/S0166432810004158>
- 15) Hofstetter Sebastian: Überblick kognitiver Architekturen [Bericht]. - Technische Universität
München : Florian Röhrbein, 2012.

8 Abbildungsverzeichnis

Abbildung 1: Überblick über Wissensarten.....	5
Abbildung 2: Cowan-Oberauer modell ¹⁴	6
Abbildung 3: Aktivierungsausbreitung mit Assoziationsmechanismen ¹¹	7
Abbildung 4: informationsverarbeitungsprozess ²	7
Abbildung 5: Architektur von SOAR ²	10
Abbildung 6: Repräsentation eines Zustandes ⁶	11
Abbildung 7: repräsentation eines Zustandes ⁶	11
Abbildung 8: Entscheidungszyklus ⁶	11
Abbildung 9: Hauptmodule von ICARUS.....	12