

## **Brauchen Interface Agenten Emotionen?<sup>1</sup>**

Nicole C. Krämer & Gary Bente  
Universität zu Köln, Psychologisches Institut

### **Zusammenfassung**

Da anthropomorphe Interface Agenten neben reiner Informationsvermittlung im Sinne eines benutzerfreundlichen Ansatzes auch sozial-kommunikative bzw. sozio-emotionale Aufgaben erfüllen sollen, beginnen zahlreiche Forscher ihre Implementationen mit der Modellierung der Emotionen des Agenten. Durch den Rückgriff auf Emotionsmodelle soll sichergestellt werden, dass in Abhängigkeit von Nutzerverhalten und Situation ein spezifischer interner Systemzustand entsteht. Dieser interne Zustand wird (meist nonverbal) gezeigt und soll den Benutzer in eine gewünschte Richtung beeinflussen. Anhand von sowohl grundlagenwissenschaftlichen als auch angewandten Theorien und Modellen soll deutlich gemacht werden, inwieweit Architekturen, die ohne Emotionsmodellierung auskommen, einfacher zum Ziel führen können. Eine erste auf solchen Überlegungen beruhende Implementation im Rahmen der Schnittstelle des Privathaushalt-Demonstrators des Projektes EMBASSI (Elektronische Multimodale Bedien- und Serviceassistenz) wird vorgestellt.

### **1 Funktion von anthropomorphen Interface Agenten**

Anthropomorphe Interface Agenten sollen zukünftig nicht nur natürlichsprachigen Dialog ermöglichen, sondern auch ein benutzerfreundliches Klima erzeugen. Diese Anforderung bezieht sich bei Anwendungen im Haushalt oder in beruflichen Settings etwa darauf, dass der Benutzer bei Unaufmerksamkeit oder Langeweile aktiviert oder bei Ärger beruhigt wird, während im Rahmen von Lernprogrammen angestrebt wird, den Benutzer zu motivieren. Das Mittel der Wahl, um den Benutzer entsprechend zu beeinflussen, ist weniger das verbale als vielmehr das nonverbale Verhalten des Agenten. In einigen Ansätzen und Architekturen werden daher interne Systemzustände implementiert, die auf Emotionskategorien rekurren (wie „Freude“ über ein richtig eingegebenes Ergebnis). An den „emotionalen“ Zustand werden spezifische nonverbale Verhaltensweisen gekoppelt, so dass die nonverbalen Signale dann auftreten, wenn ein bestimmter interner Zustand ausgelöst wurde. Dahinter steht die auch in der Psychologie weit verbreitete Annahme, dass Emotionen fest mit bestimmten nonverbalen Verhaltensweisen verbunden sind – in dem Sinne, dass Emotionen die Voraussetzung für emotionales nonverbales Verhalten darstellen sowie durch das Verhalten automatisch und unwillkürlich „ausgedrückt“ werden (Ekman 1997; Rimé 1983). Im folgenden Text soll nicht nur in Frage gestellt werden, inwieweit diese Annahmen zutreffen, sondern auch ob es sich bei der Modellierung von Systemzuständen unter Rückgriff auf dieses Konzept um ein für die Entwicklung von anthropomorphen Interface Agenten sinnvolles Vorgehen handelt. Muss der Interface Agent tatsächlich selbst Emotionen „haben“, um angemessenes nonverbales Verhalten zu zeigen und dadurch Emotionen beim Nutzer auslösen zu können? Muss zunächst ein darunter liegender Zustand modelliert werden, um Verhalten steuern zu können (ob-

---

<sup>1</sup> Die dargestellten Arbeiten sind im Rahmen des vom BMB+F geförderten Leitprojektes EMBASSI (Elektronische Multimodale Bedien- und Serviceassistenz; BMB+F Förderkennzeichen 01 IL 904 L) entstanden.

wohl dann zunächst umfangreiches Wissen über Prozesse vorliegen muss, die uns noch große Rätsel aufgeben) oder kann ein direkter Weg beschritten werden?

Dabei soll es nicht darum gehen, ob Agenten ihre „Emotionen“ auch tatsächlich fühlen sollen – diesbezüglich herrscht Einigkeit darüber, dass diese Art von Belebung nicht erwünscht und ziel führend ist. Fokussiert wird statt dessen, ob man Annahmen über das Funktionieren von Emotionen nutzen sollte, um das Verhalten des Interface Agenten zu steuern, wenn das eigentliche Ziel die Beeinflussung emotionaler Reaktionen beim Benutzer darstellt.

## 2 Stand der Forschung zur Emotionsmodellierung

Um die gestellten Fragen beantworten zu können, soll im Folgenden zunächst aufgeführt werden, welche Modelle und Architekturen vorgestellt wurden. Im Anschluss wird eine kommunikationspsychologischen Kontroverse dargestellt, die die Annahme des Zusammenhangs von Emotion und nonverbalem Verhalten zumindest in Frage stellt.

### 2.1 Modelle und Architekturen

Rosalind Picard (1999) vertritt unter dem Begriff „affective computing“ die These, dass nicht nur die Emotionen des Nutzers im Rahmen des Mensch-Maschine Dialoges erkannt und genutzt werden sollten, sondern auch der Computer Emotionen zeigen sollte (Picard 1999; Picard & Cosier 1997). Hintergrund ist die Annahme, dass durch Emotionen vor allem auch kognitive Prozesse gefördert werden sowie Kommunikation erleichtert wird. Computer werden mit Emotionsmodellen ausgestattet, damit sie die menschlichen Fähigkeiten besser simulieren können und über die Voraussetzung zum empathischen Mitfühlen verfügen – mit dem Ziel, den Computer zu einem persönlichen Begleiter zu machen, der auf die Stimmung seines Besitzers eingeht und seine Bedürfnisse zu erfüllen sucht.

Insbesondere im Bereich von pädagogischen Programmen wird ferner angenommen, dass anthropomorphe Interface Agenten vor allem dann Vorteile bringen, wenn sie emotionale Verhaltensweisen zeigen, durch die sie den Nutzer – und in diesem Falle insbesondere Kinder – motivieren: „Drawing on a rich repertoire of emotive behaviors to exhibit contextually appropriate facial expressions and expressive gestures, they [pedagogical agents] can exploit the visual channel to advise, encourage, and empathize with students“ (Lester et al. 2000 S. 124). Laut Elliott, Rickel und Lester (1999) kann der Lernende durch das Zeigen von Reaktionen über den Fortschritt ange regert werden, sich selbst mehr für seinen Fortschritt zu interessieren. Darüber hinaus könne der Agent intervenieren, wenn der Lernende frustriert zu werden scheint oder er könne Enthusiasmus für die behandelte Materie zeigen und im Lernenden hervorrufen. Auf Basis einer Architektur, die Lester et al. (2000) als „emotive-kinesthetic behavior framework“ bezeichnen, werden auf Grundlage der jeweiligen Situation „emotive categories“ ausgewählt, die die Verbindung zwischen den jeweiligen Zielen der Konversation und dem Verhalten darstellen. Angelehnt wurde das Vorgehen an den Affective Reasoner (AR) von Elliott (1992), der auf das Emotionsmodell von Ortony, Clore und Collins (1988) zurückgreift. Sowohl Elliott (1992) als auch Lester et al. (2000) rekurren auf ähnliche Grundannahmen: „The emotive-kinesthetic behavior sequencing framework exploits the fundamental intuition behind AR – namely, that the emotive states and communication are intimately interrelated“ (Lester et al. 2000, S. 142). Auf Basis eines einfachen one-to-one-mapping Modells werden verschiedenen Situationen und Sprechakten emotionale Zustände zugeordnet, die wiederum mit bestimmten Verhaltenweisen einhergehen. Zum Repertoire gehören *happiness, elation, sadness, fear, excitement, envy, shame, gloating*, die mit Hilfe spezifischer

anatomischer Emotionsträger wie Hände, Augen, Augenbrauen, Mund, Kopfneigung, Haltung und Gestik ausgedrückt werden.

Moldt und von Scheve (2001) stellen ein soziologisch orientiertes Emotionsmodell vor, mit dessen Hilfe sie emotionale Interface Agenten gestalten möchten, die nicht nur in dyadischer Interaktion, sondern auch in sozialen Aggregaten einsetzbar sind (vgl. Moldt & von Scheve 2002). Die Autoren argumentieren, dass unerfahrenen Nutzern ein Zugang nur dann gewährt wird, wenn innerhalb der Interaktion auf bekannte Muster rekurriert wird. Dieses durchaus überzeugende Argument führt allerdings nicht zwangsläufig dazu, dass die Modellierung von Emotionen nach dem Vorbild der Emotionszusammenhänge beim Menschen nahe liegt – wie die Autoren weiter ausführen. Im Gegensatz zu dem Ansatz von Lester et al. (2000) vertreten Moldt und von Scheve zwar keine Modellierung interner Emotionszustände, doch sie fördern Emotionsregulation mit Hilfe eines „rule-based system of emotion work“ (Moldt & von Scheve 2001, S. 292). Auch hierbei wird auf Emotionsmodelle rekurriert, allerdings auf solche mit symbolisch-interaktionistischer Provenienz unter Einbezug sozialer Normen: „Agents must have knowledge about feeling rules, correspondence rules, coding rules, manifestation rules and the interrelations between these rules“ (S. 292). Dennoch vertreten Moldt und von Scheve vor allem eine kommunikationsorientierte Sichtweise, erkennbar etwa an der genutzten Terminologie: So wollen sie ein „soziales Selbst“ implementieren.

Auch im Rahmen der Robotik wird z.T. die Verwendung von Emotionskategorien zur Modellierung von Systemzuständen als sinnvoll angesehen (vgl. Dautenhahn & Christaller 1997). Das unmittelbar verfolgte Ziel ist hier allerdings ein anderes: Die Emotionen sollen hier der Intelligenz und der Selbststeuerung dienen und werden direkt an sensomotorische Fähigkeiten gekoppelt. Sie sind nicht – wie im Falle des Interface Agenten - hauptsächlich auf die Manipulation der Emotionen und Verhaltensweisen des Nutzers ausgerichtet. Die Diskussion über Sinnhaftigkeit der Emotionsmodellierung ist somit vor einem anderen Hintergrund zu führen und wird im Folgenden nicht weiter berücksichtigt.

## 2.2 Kontroverse zum Zusammenhang von Emotion und nonverbalem Verhalten

Die Modellierung von internen Systemzuständen nach dem Beispiel der Emotionen kann im Rahmen der Entwicklung anthropomorpher Interface Agenten nur hilfreich sein, wenn man über eindeutige Regeln darüber verfügt, a) welcher Input - im Sinne von Nutzerverhaltensweisen - welchen „emotionalen“ Systemzustand auslöst und b) durch welche Verhaltensweisen die Emotionen ausgedrückt werden. Hinzu kommt, dass man ebenfalls Annahmen darüber benötigt, welche Attributionen und sozio-emotionalen Reaktionen im Nutzer ausgelöst werden, wenn diese die Verhaltensweisen des Agenten beobachten. Während empirische Ergebnisse zum letztgenannten Aspekt auch benötigt werden, wenn man die Implementierung der im Weiteren aufzuführenden alternativen Architektur anstrebt, könnten die als a und b gekennzeichneten Fragen in diesem Fall unberücksichtigt bleiben. Dies ist insofern vorteilhaft, als nicht nur zur Frage, welche Situationen welche Emotionen hervorrufen (siehe a), wenig definitive Erkenntnisse verfügbar sind, sondern vor allem offen ist, in welcher Beziehung Emotionen und die sie „ausdrückenden“ Verhaltensweisen stehen (siehe b). Eine entsprechende, seit Jahren bestehende Kontroverse zum Zusammenhang von Emotion und nonverbalem Verhalten soll im Folgenden vorgestellt werden.

Diskutiert wird, ob das nonverbale Verhalten durch emotionale Zustände entsteht und die Vermittlung einer Botschaft eher sekundären Charakter hat oder ob vielmehr nonverbales Verhalten ausschließlich kommunikativ und sozial motiviert ist. Vertreter der erstgenannten Position, die als „classical emotions view“ oder „readout“-Ansatz bezeichnet wird, gehen unter Rückbezug auf

Darwin (1872) von einer ausdrucks- bzw. emotionspsychologischen Perspektive aus: Jegliche, durch unterschiedliche Stimuli induzierte Emotionen rufen laut der Vertreter unwillkürlich non-verbales oder mimisches Verhalten hervor. Die entstehenden mimischen Muster seien spezifisch für die jeweilige Emotion, die anhand dessen erkannt werden kann (Ekman 1997; Tomkins 1962). Vertreter der davon abweichenden Position dagegen bestreiten den unmittelbaren Zusammenhang zwischen nonverbalem Verhalten und Emotionen und betonen statt dessen den kommunikativen Aspekt (Chovil 1991; Frey 1999; Fridlund 1991). Mimik, Gestik und Körperhaltung werden nicht als emotional, sondern als ausschließlich sozial bedingt angesehen. Ausgangspunkt für diese Annahmen stellen verschiedene empirische Ergebnisse dar, die zeigen, dass das nonverbale Verhalten keineswegs ein automatischer Ausdruck von emotionalen Zuständen ist, sondern der Sinn und Zweck vielmehr ausschließlich in der Kommunikation liegt. Kraut und Johnston (1979) etwa weisen am Beispiel des Lächelns nach, dass nicht Freude oder Glück mimische Reaktionen hervorrufen, sondern soziale Situationen. So lächeln Bowlingspieler nicht etwa, wenn sie den Erfolg ihres Wurfes beobachten, sondern wenn sie sich anschließend zu ihren Mitspielern umdrehen.

Das elaborierteste Konzept des sozial-kommunikativen Ansatzes stammt von Fridlund (1991): Anhand von Beispielen und unter Rückbezug auf empirische Ergebnisse verdeutlicht er, dass im Laufe der Evolution nicht diejenigen begünstigt sind, die ihre Gefühle und Absichten unmittelbar zeigen, sondern vielmehr diejenigen, die in der Lage sind, mimische und nonverbale Verhaltensweisen im kommunikativen Kontakt sinnbringend und manipulativ einzusetzen. Funktion der mimischen Displays sei es nicht, Basisemotionen ablesbar zu machen, sondern soziale Begegnungen zu organisieren. Es wird postuliert, dass ein unmittelbarer und unbewusst wirkender Zusammenhang zwischen Emotion und nonverbalem Verhalten gar nicht besteht und jegliches Verhalten nicht durch interne Zustände, sondern durch soziale Ziele gesteuert wird und somit immer manipulativ (in einem positiven Sinne) ist. Zusammenfassend lässt sich festhalten, dass die Theorie von Fridlund (1991) nahe legt, dass Emotionen und nonverbales Verhalten nicht notwendigerweise verbunden sind, sondern nonverbales Verhalten – losgelöst von etwaig vorherrschenden Emotionen – ganz auf den Rezipienten ausgerichtet ist. Daraus lässt sich für die Steuerung der nonverbalen Signale bei anthropomorphen Interface Agenten ableiten, dass eine Modellierung der Emotionen nicht nur nicht notwendig, sondern eventuell auch wenig hilfreich ist, da diese nicht eins zu eins mit bestimmten Verhaltensweisen verbunden werden können. Statt dessen lässt sich auf Basis der Annahme, dass nonverbale Verhaltensweisen unabhängig von Emotionen zur Beziehungsregulation eingesetzt werden (vgl. auch Frey 1999), die Empfehlung ableiten, die Steuerung nonverbaler Signale mehr von der intendierten Wirkung auf den Nutzer abhängig zu machen.

### 3 Vorschlag für eine alternative Architektur

Wird bereits im Rahmen der Grundlagenforschung zur Kommunikation kontrovers diskutiert, inwieweit ein direkter Zusammenhang zwischen Emotion und nonverbalem Ausdruck besteht (Ekman 1997; Fridlund 1991; vgl. Krämer 2001), so ist eine Implementation dieses ungeklärten Zusammenhangs im Rahmen von anthropomorphen Agenten ein Umweg, der bei genauerer Betrachtung einiges an Sinnhaftigkeit entbehrt (vgl. Bente & Otto 1996). Vielmehr bietet sich im Rahmen von anthropomorphen Interface Agenten eine unmittelbare Orientierung an der Wirkung auf den Rezipienten an. Um die Frage zu beantworten, welches Verhalten gezeigt werden sollte, um eine spezifische Reaktion auf Seiten des Nutzers zu erzielen, muss nicht etwa gefragt werden, welcher interne Systemzustand im Agenten zu modellieren ist. Vielmehr kann die Antwort auf diese Frage aus Erkenntnissen über spezifische sozio-emotionale Wirkungen nonverbaler Verhaltensweisen hergeleitet werden: Ein spezifisches Verhalten wird immer dann eingesetzt, wenn es auf der Basis des „Wissens“ zur Wirkung effektiv zu sein verspricht. Es muss somit geprüft wer-

den, ob bisherige empirische Ergebnisse zur Wirkung nonverbaler Signale aussagekräftig genug sind, um ein Regelwerk abzuleiten.

Auf Basis von Abbildung 1 und 2 lassen sich die unterschiedlichen Herangehensweisen vergleichen: Während in Abbildung 1 ein Modell über interne Zustände gebildet wird, ist die Struktur in Abbildung 2 wesentlich handlungsbezogener und direkter auf die beim Nutzer angestrebte Reaktion bezogen. Im Rahmen der hier vorzuschlagenden Architektur wird die Modellierung des Zusammenhangs zwischen der Situation, der „internen“ Reaktion des Agenten und den sich daraus ergebenden Verhaltensweisen hinfällig. Statt dessen tritt der Zusammenhang zwischen Verhalten des Agenten und dem Auslösen bestimmter Reaktionen und Emotionen im Benutzer in den Vordergrund. Ist nun bekannt, welches Verhalten welche Reaktionen auslöst, kann dieses Wissen genutzt werden, um in Situationen, in denen man entscheiden kann, welches Verhalten oder Erleben des Benutzers man erreichen möchte, ein entsprechendes Verhalten des Agenten vorzusehen.

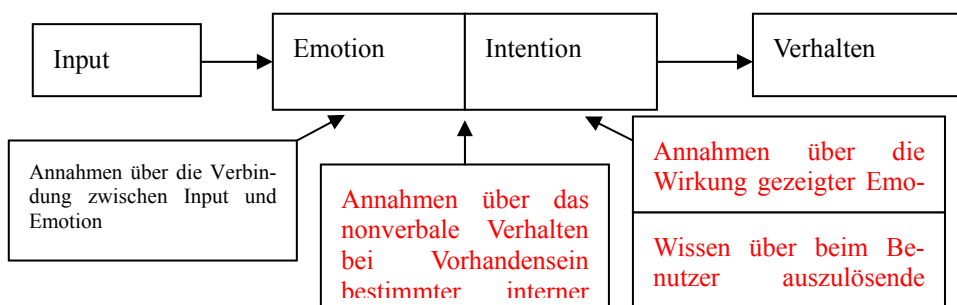


Abb. 1: Skizzierung gängiger Implementierungen unter Rückgriff auf die Modellierung zugrunde liegender Zustände

Wenn der Agent etwa verkünden muss, dass die Aufnahme des Lieblingsfilmes nicht erfolgreich war, soll er durch entsprechende nonverbale Begleitverhaltensweisen erreichen, dass der Benutzer nicht wutentbrannt das Gerät umtauscht, sondern sich beruhigt. Dazu muss allerdings nicht der Agent „traurig“ werden – geleitet von der Hoffnung, dass wenn dieser traurig ist, der Benutzer sieht, dass er den Fehler bedauert etc. - sondern es reicht aus, den Agenten in solchen Fällen ein Verhalten zeigen zu lassen, von dem nachgewiesen wurde, dass es eine besänftigende Wirkung in bestimmten Situationen hat. Dies kann im Ergebnis natürlich durchaus ein Verhalten sein, das bestimmte Emotionszuschreibungen durch den Benutzer hervorruft – aber ohne dass es notwendig war, den Zusammenhang von Situation, Emotion und Verhalten zu modellieren und den Agenten in einen „emotionalen Zustand“ zu versetzen.

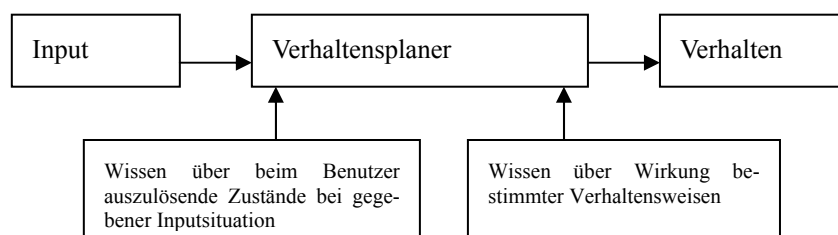


Abb. 2: Alternative Architektur unter Rückgriff auf ein Regelwerk zur Wirkung nonverbaler Verhaltensweisen

Solange Interaktionen mit Agenten noch so wenig komplex sind, wie dies momentan der Fall ist, ergibt sich bezüglich der Effizienz der beiden Ansätze vielleicht eventuell noch gar kein Unter-

schied. Tatsächlich werden momentan höchstens extrem reduzierte Dialoge geführt, bei denen an vordefinierten Stellen „Emotionen“ gezeigt werden. Wenn aber in Zukunft Dialoge flexibler geführt werden, wird die Notwendigkeit, stimmige und empirisch gesicherte Regelwerke zu implementieren, die echtzeitfähig Reaktionen steuern, zu einer Bevorzugung nicht nur von Modellen, die mehr nutzbare empirische Befunde hervorgebracht haben, sondern auch einfacheren Modellen führen. Hier wird sich in den nächsten Jahren erweisen müssen, welche empirischen Erkenntnisse besser verfügbar sind und effektiver implementiert werden können – die über den Zusammenhang von Input, Emotionen und Verhaltensweisen oder die über die sozio-emotionale Wirkung von bestimmten Signalen.

#### 4 Implementation im Rahmen des EMBASSI-Demonstrators

Im Rahmen des Projektes EMBASSI wurde ein Demonstrator für den Bereich Privathaushalt, der insbesondere Auswahl, Aufnahme und Rezeption von TV und Video unterstützt, entwickelt, dessen Schnittstelle einen anthropomorphen Interface Agenten enthält (vgl. <http://www.embassi.de>). Durch unterschiedliche beteiligte Partner wurden ein Dialogmanagementsystem, ein Modul zur Sprachgeneration, ein Modul zur Planung des multimodalen Out- und Inputs, eine Benutzeroberfläche, ein Tool zur Sprachsynthese sowie ein anthropomorpher Charakter entwickelt. Damit die anthropomorphe Figur Bewegungen und Mimik zeigen kann, die den wiederzugebenden Inhalten entsprechen (ein nicht triviales Problem, da das den Agenten steuernde Modul zunächst einmal nicht weiß, was der Agent äußert), müssen Regeln vorliegen, die Auskunft darüber geben, welche Art von Aussage durch welche nonverbalen Signale unterstützt werden sollen, um einen angestrebten Zustand im Benutzer auszulösen. Als Ausgangspunkt dazu dienen die im Dialogmanager erzeugten Aussagetypen, die an der Sprechakttheorie (vgl. Searle, 1969) orientiert sind und bereits Auskunft über die Semantik der zu formulierenden Aussage geben. Tabelle 1 gibt einen Überblick über die auf empirischer Grundlage (vgl. Krämer, 2001) entwickelten Regeln, die den im Rahmen des Demonstrators vorkommenden Sprechakten zugeordnet wurden.

Tabelle 1: Übersicht über Sprechakte, verfolgte Intentionen und zu zeigendes Verhalten

| Sprechakt   | Intention/zu erreichendes beim Benutzer                | Konkretes Verhalten Agent  | Zeitaspekt  |
|---|--|--|---|
| <b>message_greeting:</b> Das System begrüßt den Benutzer, am Anfang der Interaktion | Der Benutzer fühlt sich freundlich und höflich begrüßt | Augenbrauen hoch und Augen 1 Sekunde größer<br>Lächeln<br>Winken (optional)<br>Lächeln<br>Kopf schräglegen (ca. 10 Grad nach rechts) | Beginn der Einblendung<br><br>Vor Beginn<br>Vor Beginn<br>Nach Begrüßung<br>Nach Begrüßung (für 4 Sekunden) |
| <b>message_closing:</b> Wird eingesetzt am Ende der Interaktion mit dem Benutzer.   | Der Benutzer fühlt sich freundlich verabschiedet       | Lächeln  | Nach Verabschiedung   |
| <b>message_inform</b> [status: <b>warning</b> ]                                     | Benutzer soll Dringlichkeit wahrnehmen                 | Augenbrauen anheben, Augen größer<br>Hände anheben, Handflächen zum Benutzer (evtl. nur eine Hand heben mit Zeigefinger nach oben)   | Gesamtzeit der Äußerung<br>Zu Beginn, dann langsam runternehmen   |
| <b>message_inform</b> [status: <b>busy</b> ]  | Benutzer soll geduldig bleiben                         | Kopf geringfügig nach unten neigen   | Gesamtzeit der Äußerung   |

|   |   |  |  |
|---|---|--|--|
|   |   | Wegschauen (turn bleibt bei System)  | Nach Ende Äußerung   |
| <b>message_inform</b> [status: <b>error</b> ]   | Benutzer soll besänftigt werden                                   | Lächeln<br>Kopf schräglegen (ca. 10 Grad nach rechts)  | Nach Ende Äußerung<br>Nach Ende Äußerung (für 4 Sekunden)  |
| <b>message_inform</b> [status: <b>ok</b> ]  | Positive Gefühle sollen induziert werden                          | Leicht ausgeprägtes Lächeln  | Nach Ende der Äußerung   |
| <b>message_inform</b> [status: <b>failed</b> ]  | Benutzer soll besänftigt werden                                   | Schultern anheben, Hände (mit Handflächen nach oben) bis Ellbogenhöhe<br>Lächeln<br>Kopf schräglegen (ca. 10 Grad nach rechts)                             | Beginn der Äußerung, dann langsam senken<br><br>Nach Ende Äußerung<br>Nach Ende Äußerung (für 4 Sekunden)            |
| <b>message_inform</b> [status: <b>offer</b> ]<br>z.B. proaktives Angebot                                  | Benutzer soll Interesse verspüren                                 | Zeigefinger heben<br><br>Leicht ausgeprägtes Lächeln   | Zu Beginn, nach 3 Sek. langsam senken<br>Gesamtzeit; evtl. nach Ende verstärken                                      |
| <b>message_accept</b> : Wird für Zustimmungen benutzt.  | Benutzer soll sich bestätigt fühlen                               | Nicken   | Vor Beginn der Äußerung  |
| <b>message_reject</b> : Wird für Verneinungen benutzt, wobei eine Tatsache als falsch bezeichnet wird.    | Ein Fehler soll vermittelt werden, ohne dass Ärger ausgelöst wird | Kopf leicht schütteln, ansonsten neutral   | Vor Beginn der Äußerung  |
| <b>message_command</b> : Kann für Vorschläge bzw. dringlichen Hinweise an den Benutzer eingesetzt werden. | Benutzer soll Dringlichkeit wahrnehmen                            | Augenbrauen anheben, Augen größer<br>Hände mit Handflächen zum Benutzer anheben (evtl. nur eine Hand heben mit Zeigefinger)                                | Gesamtzeit der Äußerung<br>Zu Beginn, dann langsam runternehmen  |
| <b>message_cancel</b> : plötzliche turn-Übergabe nach Unterbrechung durch Nutzer                          | Nutzer soll merken, dass er sprechen darf                         | Nutzer ansehen   | Mindestens bis Nutzer spricht  |
| <b>message_acknowledge</b> : Bestätigt die erfolgreiche Durchführung eines Auftrages.                     | Positive Gefühle induzieren                                       | Lächeln<br><br>Evtl. emblematische Gesten wie Daumen heben   | Kurz vor Äußerung beginnen, während Äußerung halb, nach Ende verstärken  |
| <b>message_correct</b> : Korrigiert eine falsche Annahme des Benutzers.                                   | Korrektur, ohne dass Ärger ausgelöst wird                         | Neutral  | Gesamtzeit   |
| <b>query_input</b> : wenn das System irgendwelche Eingaben vom Benutzer erwartet.                         | Benutzer soll sich aufgefordert fühlen                            | Blick zum Benutzer<br>Kopf leicht schräglegen (ca. 5 Grad nach links)  | Mit Ende der Äußerung 3 Sek. nach Ende der Äußerung, für 5 Sek.  |
| <b>query_y/n</b> : Erzeugt eine Erwartung auf eine einfache <i>Ja-</i> bzw. <i>Nein-</i> Antwort.         | Benutzer soll sich aufgefordert fühlen                            | Blick zum Benutzer<br>Kopf leicht schräglegen (ca. 5 Grad nach links)  | Mit Ende der Äußerung 3 Sek. nach Ende der Äußerung, für 5 Sek.  |
| <b>query_selection</b> : Impliziert eine Auflistung von Alternativen                                      | Benutzer soll das Verständnis erleichtert werden                  | Gestik, beat gesture (mal rechts vom Körper, mal links)<br>Alternativ: Zeigegeste<br>Blick zum Benutzer<br>Kopf leicht schräglegen (ca. 5 Grad nach links) | Jede Alternative durch beat gesture markieren<br><br>Mit Ende der Äußerung 3 Sek. nach Ende der Äußerung, für 5 Sek. |
| <b>query_request acknowledge</b> : Wird bei Unsicherheit während  | Benutzer soll sich aufgefordert fühlen                            | Blick zum Benutzer   | Mit Ende der Äußerung  |

|  |  |  |   |
|--|--|--|---|
| der Sprachanalyse eingesetzt (z.B. fehlerhafte Spracherkennung)  |  | Kopf leicht schräglegen (ca. 5 Grad nach links)  | 3 Sekunden nach Ende der Äußerung, für 5 Sekunden   |
| <b>Query_request_repair:</b> Ist ebenfalls für falsche Erkennungsergebnisse bzw. Korrekturen, die vom Benutzer angeführt wurden, gedacht. Reparatur-Strategien sind Wiederholungsanforderungen an den Benutzer oder starre Ja/Nein-Fragen. | Benutzer soll besänftigt werden, sich dann aufgefordert fühlen | 1. Teil (nicht verstanden)<br>Schultern anheben, Hände (mit Handflächen nach oben), auf Ellbogenhöhe<br>Lächeln<br>Kopf schräglegen (ca. 10 Grad nach rechts)<br><br>2. Teil:<br>Blick zum Benutzer<br>Kopf leicht schräglegen (ca. 5 Grad nach links) | Beginn der Äußerung, dann langsam senken<br><br>Nach Ende Äußerung<br>Nach Ende Äußerung (für 4 Sekunden)<br><br>Mit Ende der Äußerung<br>3 Sek. nach Ende der Äußerung, für 5 Sek. |

Die meisten der Zuordnungen von Verhaltensweisen gehen zurück auf empirische Ergebnisse, die in Krämer (2001) zusammengestellt wurden. So basiert der Einsatz des Kopf-Schräglegens (der sogenannten Lateralflexion) auf Ergebnissen, dass dieses als positiv empfunden wird und besänftigend wirken kann. Beim Versuch der Zusammenstellung der Regeln zeigt sich jedoch auch, dass noch nicht genügend empirische Ergebnisse zur Wirkung bestimmter Verhaltensweisen (und auch deren zeitlicher Strukturierung) verfügbar sind. Daher musste an zahlreichen Stellen auf emblematische Gesten wie z.B. dem Anheben der Hände und Schultern zur Vermittlung, dass der Nutzer Informationen geben muss, zurückgegriffen werden. Hierbei handelt es sich weniger um tatsächliche nonverbale Verhaltensweisen – geschweige denn um momentan im Zentrum der Forschungsaufmerksamkeit stehende subtile Dynamiken - sondern vielmehr um kulturell vermittelte, in ihrer Bedeutung feststehende Zeichen. Im Rahmen weiterer Forschung – die auch von der Nutzung der Interface Agenten profitieren kann (vgl. Bente & Krämer 2002) – müssen somit weitere Ergebnisse zur Wirkung nonverbalen Verhaltens generiert werden.

## 5 Fazit

Die Ausführungen sollten zeigen, dass der Umweg über das Nachbilden der beim Menschen angenommenen Emotionszusammenhänge zumindest im Falle der Interface Agenten gar nicht gegangen werden muss. Es wurde vorgeschlagen, in Zukunft Vorgehensweisen zu erproben, die „straightforward“ das Wissen über die Wirkung nonverbalen Verhaltens nutzen, um ein benutzerfreundliches Klima auf einfacherem Wege herzustellen als durch die Modellierung von Emotionen. Um auf dieser Basis eine erfolgreiche Implementation allerdings überhaupt leisten zu können, müssen mehr Erkenntnisse über die Wirkung – insbesondere subtiler – nonverbaler Verhaltensweisen generiert werden, die die Regelwerke entsprechend erweitern. Problematisch bleibt allerdings, dass die Wirkung der Signale immer kontextabhängig ist (vgl. Krämer 2001) und von Variablen wie Situation, Erscheinungsbild des Senders und sonstigen Verhaltensweisen des sendenden Agenten beeinflusst wird. Letztendlich wird nicht endgültig vorhergesagt werden können, welche (emotionalen) Reaktionen ein spezifisches Verhalten des Interface Agenten beim Nutzer auslösen wird – dieses Problem löst aber auch der Rückgriff auf Emotionen nicht. So wie aber auch der Mensch lernt, Verhalten zu zeigen, mit dem er das Gegenüber beeinflussen bzw. ein positives Klima herstellen kann, sollte auch die explizite Zusammenstellung dieser Regeln mit der Zeit gelingen. Durch die alternative Herangehensweise wird der Blick – auch und vor allem in bezug auf die Forschung – in eine neue Richtung gelenkt, die auf Dauer vielversprechende Umsetzungen ermöglichen sollte.



## Literatur

- Bente, G. & Krämer, N. C. (2002). Virtuelle Gesten: VR-Einsatz in der nonverbalen Kommunikationsforschung. In: G. Bente, N. C. Krämer & A. Petersen (Hrsg.), *Virtuelle Realitäten* (S. 81-107). Göttingen: Hogrefe.
- Bente, G. & Otto, I. (1996). Virtuelle Realität und parasoziale Interaktion. In: *Medienpsychologie: Zeitschrift für Individual und Massenkommunikation*, 8, 217-242.
- Chovil, N. (1991). Social determinants of facial displays. In: *Journal of Nonverbal Behavior*, 15 (3), 141-154.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: J. Murray.
- Dautenhahn, K. & Christaller, T. (1997). Remembering, rehearsal and empathy - towards a social and embodied cognitive psychology for artefacts. In: S. O'Nuallain, P. McKeivitt & E. MacAogain (Eds.), *Two Sciences of Mind*. Amsterdam: John Benjamins.
- Ekman, P. (1997). Expression or communication about emotion. In: N. L. Segal & G. E. Weisfeld (Eds.), *Uniting psychology and biology: Integrative perspectives on human development* (pp. 315-338). Washington: American Psych. Association.
- Elliott, C. (1992). *The affective reasoner: A process model of emotions in a multi-agent system*. Ph.D thesis, Institute for the Learning Sciences, Northwestern University.
- Elliott, C. Rickel, J. & Lester, J. (1999). Lifelike pedagogical agents and affective computing. An exploratory synthesis. In: M. Woolridge & M. Veloso (Eds.), *Artificial intelligence today* (pp. 195-212). Berlin: Springer.
- Frey, S. (1999). *Die Macht des Bildes*. Bern: Huber.
- Fridlund, A. J. (1991). Evolution and facial action in reflex, social motive, and paralanguage. In: *Biological Psychology*, 32 (1), 3-100.
- Krämer, N. (2001). *Bewegende Bewegung. Sozio-emotionale Wirkungen nonverbalen Verhaltens und deren experimentelle Untersuchung mittels Computeranimation*. Lengerich: Pabst.
- Kraut, R. E. & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. In: *Journal of Personality and Social Psychology*, 37 (9), 1539-1553.
- Lester, J. C., Towns, S. G., Callaway, C. B., Voerman, J. L. & FitzGerald, P. J. (2000). Deictic and emotive communication in animated pedagogical agents. In: J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), *Embodied Conversational agents* (pp. 123-154). Boston: MIT Press.
- Moldt, D. & von Scheve, C. (2001). Emotions and multimodal interface-agents: A sociological view. In: H. Oberquelle, R. Oppermann & J. Krause (Hrsg.), *Mensch und Computer 2001* (S. 287-295). Stuttgart: Teubner.
- Moldt, D. & von Scheve, C. (2002). Emotions in hybridsocial aggregates. In: M. Herczeg, W. Prinz & H. Oberquelle (Hrsg.), *Mensch und Computer 2002* (S. 343-352). Stuttgart: Teubner.
- Ortony, A., Clore, G. L. & Collins, A. (1988). *The cognitive structure of emotion*. New York: Cambridge University Press.
- Picard, R. W. & Cosier, G. (1997). Affective Intelligence – The missing link. In: *BT Technology*, 14 (4), 150-161.

- Picard, R. (1999). *Affective Computing*. Cambridge: MIT Press.
- Rimé, B. (1983). Nonverbal communication or nonverbal behavior? Towards a cognitive-motor theory of nonverbal behavior. In: W. Doise & S. Moscovici (Eds.), *Current issues in European social psychology* (pp. 85-141). Cambridge: Cambridge University Press.
- Searle, J. R. (1969). *Speech acts*. Cambridge: Cambridge University Press.
- Tomkins, S. S. (1962). *Affect, imagery, consciousness. The positive affects*. New York: Springer.

## Adressen der Autoren

Dr. Nicole Krämer / Prof. Dr. Gary Bente  
Universität zu Köln  
Psychologisches Institut  
Bernhard-Feilchenfeld-Str. 11  
50969 Köln  
[nicole.kraemer@uni-koeln.de](mailto:nicole.kraemer@uni-koeln.de)/[bente@uni-koeln.de](mailto:bente@uni-koeln.de)