日本における人工知能研究をめぐる

オーラルヒストリーVol.5（斉藤　康己）

（AIR 温故知新プロジェクト：オーラル・ヒストリーシリーズ）

AIR: Acceptable Intelligence with Responsibility

2019 年 7 月

AIR 温故知新プロジェクト：オーラル・ヒストリーシリーズについて

人工知能研究は情報学に軸足を置きながらも、その発足当時より、人文社会科学をはじめとした他領域とのコラボレーションによって発展してきました。そうしたコラボレーションは、第 1 次人工知能ブームから現在に至るまで、様々な形態・ネットワークで、時に拡大・縮小を繰り返しながら展開してきました。

2016 年現在、第 3 次人工知能ブームとも言うべき状況にあります。こうした状況下において、人工知能研究者やそれに関連する仕事をしている人びとは、産業への応用可能性を基軸とした民間・政府資金の集中投下といった政策的展開や、それに伴う組織の拡大・再編成、そして多様な研究・応用ネットワークの形成など、目まぐるしい状況変化の只中に身を置いていると言えます。そのようななかで、どのような針路をとっていくべきなのか、問題として経験されることは何か、それはどのように解決すべきなのかといったことを考えることは非常に重要です。

他方で、第 3 次人工知能ブームと先に述べたように、すでに私たちの社会は第 1 次・第 2 次ブームを経験しています。つまり、すでに来た道であるならば、そこから学び、現状に対してそれを活かしていくこと――まさに温故知新――もまた必要なのではないでしょうか。私たち AIR は、こうした問題意識のもと、2016 年現在も現役で活躍している方が多い第 2 次人工知能ブームに身を置いていた先人たちへのインタビューを企画しました。

こうした取り組みは、期せずして異分野協働体制となった AIR のメンバー間の相互理解・共通言語構築への一助になることも企図しています。また、インタビューを通して、人工知能をめぐる科学技術史や学際的連携の在り方についての研究などの発展的展開も将来展望として考えています。

一方、先人たちの経験や教えを私たち AIR のみで共有するのではなく、現在第 3 次人工知能ブームと言われるなかで一般に広くそれを公開することは社会的意義があると考えます。そこで、インタビューの書き起こしデータを公開する運びとなりました。これが広く読まれ、今後の人工知能研究あるいはそれに関連するさまざまな取り組みを進める際の一助となれば幸いです。なお、本インタビューは、研究用途に限り（要出典明記）使用を認めます。

2016 年 4 月
Acceptable Intelligence with Responsibility

目次

目次・概要…… i
斉藤康己先生略年譜…… ii
インタビュー逐語録…… 1-65
附録 I AIUEO1984 年 7 月 21 日発表資料

概要

■インタビュイー（※所属はインタビュー時点（2018 年 10 月））
斉藤康己（京都大学情報環境機構 IT 企画室　教授）

■インタビュアー（※所属はインタビュー時点（2018 年 10 月））
江間有沙（東京大学）[編集]
大澤博隆（筑波大学）
神崎宣次（南山大学）
久木田水生（名古屋大学）
駒谷和範（大阪大学）
服部宏充（立命館大学）

■実施日時
2018 年 10 月 22 日

■場所
京都大学吉田キャンパス　学術情報メディアセンター北館 4 階 410 号室

■発行
AIR: Acceptable Intelligence with Responsibility

■発行年
2019 年 7 月

■本インタビューの取り扱いについて
本インタビューは研究利用に限り使用を認める。なお、使用時は必ず本報告書からの引用を明記すること。

斉藤康己先生略年譜

1953 年 8 月 20 日生
1976 年東京大学工学部計数工学科数理コース卒業。
1978 年英国エセックス大学計算機科学科修士卒業（日本政府給費留学生として）。
1979 年東京大学大学院情報工学専攻修士卒業。
1979 年日本電信電話公社武蔵野電気通信研究所基礎研究所に入所。
計算機科学、人工知能、認知科学等の研究に従事。
1984 年から 1 年間フランス INRIA 客員研究員（フランス政府給費留学生として）。
1986 年頃 NTT 研究所から米国の ARPA ネット（後にインターネットとなる）への接続の面倒をみる。
1987 年文書整形システム TeX を日本語化し、jTeX として公開（この功績により第一回 ASTEM ソフトウェア文化賞受賞）。
その後囲碁の認知科学的研究を行い 1996 年に工学博士取得（東京大学）。
1997 年から NTT 民営化後の NTT コミュニケーションズ OCN 事業部に移り OCN の販売支援、パッケージ作りなどに従事。
2000 年後半から NTT コミュニケーションズにて IPv6 プロジェクトのリーダも務める。
2005 年からは NTT ソフトウェアに移り、各種ソフトウェア開発プロジェクトの統括マネージャー、ネットワーク運用、顧客管理システムの運用、セキュリティサービスの開発などを手がけた。
2013 年 6 月から現職（京都大学情報環境機構 IT 企画室教授）。
京都大学では、全学のセキュリティ対策と学内ネットワーク KUINS の運用・保守を担当。
現在までの 5 年間に、情報セキュリティ関連規定類の改訂、全学無線 LAN インフラの整備、ソフトウェアライセンス管理の見直し、IP 電話化の検討などを行った。

著書として：
竹内郁雄編『AI 奇想曲』の第 4 章の中の「電子図書館から情報市場へ」、1992 年 1 月。
ユビキタスオフィスのテクノロジー〜ブロードバンド＆セキュリティが企業を変える〜、監修及び第 9 章執筆、電気通信協会発行、2005 年 3 月。
『アルファ碁はなぜ人間に勝てたのか』ベスト新書、2016 年 10 月。
翻訳として：
『エキスパート・システム』（フレデリック ヘイズ-ロス他著、AIUEO 共訳）、1985 年 6 月。
『メタマジック・ゲーム』（ダグラス・ホフスタッター著、共訳）、1990 年 9 月。
『リテラリーマシン　ハイパーテキスト原論』（テッド・ネルソン著、共訳）、1994 年 10 月。
など。　　斉藤先生からご提供いただきました（2018 年 9 月現在）。

服部：では、始めさせていただきます。でも、そんなにかしこまっていただく感じではないので、いつも雑談を交えながらやっていくんですけれども。よろしくお願いします。主に 80 年代ですかね、感覚的には。その頃にいろいろあったことを、今まで 4 人の先生方にお伺いしてきました。その中で、特に中島先生[1]、堀先生[2]、土屋先生[3]もそうなんですけど、AIUEO の話が結構出てきていたんですけれども、聞いていく中、言いだしっぺはっていうような話ですね（笑）。で、斉藤先生がと。いろいろ秘蔵の資料も。

斉藤：秘蔵の……（笑）。

服部：もう事前に結構、送っていただいたんですけれども、その当時のお話なんかも伺いつつ、今、思われていることなんかも併せて伺うということができればと思いますけど、よろしくお願いいたします。早速という感じなんですけれども、AIUEO って何だったのかっていうことですね。堀先生、中島先生、それぞれの視点でのお話、今まで伺ってきましたけれども、改めてというような感じになるかなと。中島先生のおっしゃるには、AI を当時、なかなか系統的に学ぶ機会はなかったっていうこともあって、あとは情報をみんなで持ち寄って、その中でいろいろと勉強をしようと。そういうような集まりだったと。何か松原さん[4]とか、橋田さん[5]のやつもあったかな。何か書いたものとかを見ると、AI をやってはならんと。まかりならんと。結構言われている中で、何かこそこそと悪いことをやるようにやっていたみたいな、何かそんな感じもあったんですけど、何か、いや、そういうものでもなく、本当に純粋に面白そうだというので、とにかくいろいろ勉強していこうというような活動だったと伺っています。そのきっかけとしては、斉藤先生がエジンバラに行かれて、それで面白いのがあるぞと。

斉藤：そこ、間違ってるんですけど。

---

[1] 中島秀之。電子技術総合研究所、産業技術総合研究所サイバーアシスト研究センター、公立はこだて未来大学、東京大学先端人工知能学教育寄付講座を経て札幌市立大学学長。人工知能、マルチエージェント。

[2] 堀浩一。国立大学協同利用機関国文学研究資料館を経て、東京大学先端科学技術研究センター。人工知能。

[3] 土屋俊。千葉大学文学部を経て、大学評価・学位授与機構教授。哲学（言語の哲学，心の哲学），高等教育質保証，認知科学，図書館情報学，音声対話。

[4] 松原仁。電子技術総合研究所を経て公立はこだて未来大学システム情報科学部教授。人工知能。

[5] 橋田浩一。1986 年より通商産業省工業技術院電子技術総合研究所。88 年から 92 年まで（財）新世代コンピュータ技術開発機構（ICOT）出向。01 年より産業技術総合研究所。13 年より東京大学大学院情報理工学系研究科。

服部：そうなんですか。

斉藤：私が行った大学はエセックス大学で、エジンバラじゃないんです。

服部：そうかそうか。エジンバラ。何でエジンバラ。

斉藤：何か前の読んだら、何カ所か間違いがあるんですけど、まあ、まあ。

　その当時、イギリスで人工知能やってたのはエジンバラと、エセックスと、サセックスぐらいだったんで、確かに間違えても不思議ではないみたいな。どういう順番で話したらいいのか分かんないですけど、もう話しだしていいんですか。

服部：はい。

斉藤：私が修士、大学の修士 1 年のときに、大学の掲示板に「交換留学制度でエセックス行けます」みたいなチラシが出たんです。工学部の 6 号館。おっ、これはちょっと面白いかなとかって思って、試験があるっていうから、試験を受けに行ったら、日本人の先生が 3 人ぐらい座ってて、最初、何か一言二言、英語で質問されたんだけど、あとは日本語で（笑）。面接で。私ともう 1 人。2 人しか受けてなくて、2 人とも合格して、2 人でエセックス行ったんです。それがエセックスへ留学する最初で、行くまでは人工知能のことはほとんど知りませんでした。計算機科学の勉強をしに行くって思って行って、計算機科学は十分、勉強しましたってか、私、一生のうちで一番勉強したのはエセックスにいる間の 1 年間だと思っていて。イギリスは、エセックス大学は 1 年間で修士が取れるコースがあるんです。で、修士を取って帰ってきたんですけど。もう何か大学のキャンパスの中にタワーっていって、学生寮があって、何階建てだったんかな。20 階建てとか、案外高い、背の高いタワーがいっぱい。いっぱい、3 つ 4 つあって、そこに住んでるんです、学生はみんな、大体。なので、もうやることないですよね、勉強しか（笑）。図書館も何か、土日もずっと開いてて、ずっと入り浸れるので、何か勉強してました、一生懸命。もうそのときの財産で、今も食ってるみたいな（笑）。その後、ほとんど勉強してないんですけど。あれです。そういうきっかけでエセックス行って。でも、その中で一番面白かったのが、やっぱり人工知能周りの、AI コリドーっていって、先生たちのいる部屋、建物の地下が AI の先生たちが何人かいるところで、そこにパット・ヘイズ[6]とか、マイケル・ブレディー[7]とか、リチャード・ボーナー[8]とか何か。当時、イギリスでは AI やってた人たちがいて、特にパット・ヘイズが面白かったんで、私はそのパットヘイズという人に付

[6] Patrick John Hayes. イギリス出身の計算機科学者、人工知能研究者。現在は、Florida の IHMC（Institute for Human and Machine Cognition）に在籍。

[7] Michael Brady. イギリスの計算機科学者。画像分析が専門。

[8] Richard Bonat? 当時 Essex 大学で計算機科学を教えていた。

いて、修士の論文、書いたんです。そこで、だから、AI の輪講みたいなのをやっていて、授業もありましたけど、授業のね。授業のね。昔の配布資料(　)[9]。これ、エセックスの、昔のカラーなんだけど、そのときの授業の資料が（笑）。　最初にエバンズ[10]。エバンズのプログラムってご存じかな、アナロジーの。知らないな、もう。今の人に聞くと、みんな、知らないんだけど、最初の頃、だから、これは初期 AI なんだと思う。第 1 次ブームのまだ延長みたいなとこだと思うんですけど、アナロジーのプログラムの話とか、有名なウィノグラード[11]の。シュルドルー[12]はご存じですよね？

服部：はい。

斉藤：ああいうのとか、GPS の話、サイモン、ニューウェル[13]の話とか、ビジョンが、Edge detection やる話とか、何かその授業、そこで受けて。授業も、まあ、面白かったですけど、とにかくパット・ヘイズの輪講のときの議論みたいなのが面白い。でも、申し訳ないけど、ほとんど覚えてないです。その中でどういう議論をしたかっていうのはほとんど覚えてないので。でも、とにかく面白かったんで、帰ってきて、帰国の報告みたいなのをする情報輪講っていうのがあって。東大、戻ってきた後ですけど、そこで「こういうのやろうよ」って言ったら、手、挙げてきたのが中島と、片桐[14]と、白井[15]っていって、まだ中京大かな。その 3 人が手を挙げてきたんです。だから、私と 4 人でやりだしたんです、AIUEO。名前が付いたの、もうちょっと後で、中島が付けたんだと、私、思ってたら、こないだ、鈴木達郎[16]っていうのが「あれは僕が付けたんだよ」って言ってましたけど、略語、ああいう略語にしたのが鈴木達郎だったのかもしれないです。最初、最初、名簿作ったのは確かに中島だったような気がするんですけど。名簿ったって、4 人（笑）。名簿もないんだけど。ちょっと数が増えてきたとき、名簿を作ってくれたのが彼で、やってたんです。やったリストが、お送りしました？　あっ、pdf で送りましたよね、これ[17]。

服部：そうですね。

斉藤：これは全然、部分的で、最初、スター

---

[9] 以下、聞き取り不可能な部分は( )で示す。

[10] Thomas G. Evans. 計算機科学者。アナロジー課題を処理する「エバンズのプログラム」を開発した。

[11] Terry Allen Winograd. 計算機科学者。SHRDLU の開発者として知られる。

[12] SHRDLU. AI 研究初期の自然言語処理プログラム。ユーザの指示にしたがって端末画面内の「積み木の世界」に存在する物体を動かすというもので、対話システムの先駆けとして知られる。

[13] Herbert Alexander Simon および Allen Newell. 「世界初の人工知能プログラム」とも呼ばれる Logic Theorist を開発した。

[14] 片桐恭弘。国際電気通信基礎技術研究所を経て、公立はこだて未来大学システム情報科学部。認知科学。

[15] 白井英俊。玉川大学工学部を経て、中京電気電子工学科教授。計算言語学、情報工学。

[16] 鈴木達郎。日本テレワーク協会客員研究員。工学院大学非常勤講師。

[17] 本書の附録 I AIUEO1984 年 7 月 21 日発表資料参照。

トから書いてあるんですけど、私が話をした1984年7月で終わってるんです。それまでの、メーリングリストでこう……。違うな。メーリングリスト、最初はないんだよな。どうやって連絡してたか、よく分かんないんですけど、4〜5 人だから、大学で会うわけなので。っていうか、前回、前の会やったら「次、誰？」って言って、もうそれだけ決めれば良かったので、彼が何持ってきて、何の論文持ってきて、何の話するかはその当日にならないと分かんないっていう感じで、ずっと続けてました。

江間：場所はずっと固定の場所だったんですか。

斉藤：いや、最初の頃は固定だったと思います。6号館だったり、あと、9号館っていうのかな。ちょっといったん門、出て。今、ないのかな。よく分かんないけど、千代田線の根津の駅のほうへ行く途中に、別のちっちゃなキャンパス。

服部：ああ、ありますね。

江間：弥生ですか。

斉藤：そうそう、弥生門を出て、ちょっと歩いたところにある、そっちでやったり、そこ、計算機センターがあったんですよ、確か。計算機センターの上のほうに部屋があったりして、そこでやったり。もう、後になってくるともう、いろんな人が来るので。違うな。やっぱり合宿かな。合宿とかだと、電総研でやったり、それからメンバーの1人、メンバーの中にお金持ちの人がいて、彼、途中で亡くなっちゃったんだけど、彼のお父さんの別荘でやらせてもらったりとか、いろいろな場所でやりました。最初は場所、決まってたような気がします。そうだ、松方君。松方財閥の御曹司だったんですよ。途中で、がんかな。すごい若くして亡くなっちゃったんです。

大澤：読んでくる論文っていうのは、もうみんな、何かもう読みたいのが勝手にあってっていう？

斉藤：勝手に。それぞれ、みんな、興味があったから、片桐なんかは言語だったし、中島は何かシステムみたいなのが多かったかな。白井君も言語だったんです、確か。だから、みんな、勝手に選んで持ってくるっていう感じでした。

江間：論文とか本って、どうやって入手されてたんですか。今はネットがあるけれど、図書館とか、最初は……。

斉藤：そうですね。いや、でも、何だろう。IJCAI[18]とかって、いつから始まったのかな。まだ始まってないかもしれない。ネタの多くは、見ると分かるかもしれないですけど、AIメモみたいな。AI メモ、大学で取ってたのかな。みんな、でも……。会議資料が多かったかもしれない。カンファレンスのプロシーディングみたいのを眺めては「あ、これ、面白そう」とか言ったりして。でも、ここに Cognitive Science（　）。

大澤：IJCAI が 69 年からで、そうですね、読まれてる論文もその辺から。

---

[18] International Joint Conference on Artificial Intelligence (IJCAI)、国際人工知能会議。

斉藤：69 年？　ああ。

駒谷：ファースト IJCAI というのがあります。

斉藤：ファースト IJCAI ぐらいからは、もう IJCAI の。AAAI[19]ができたのは随分後ですもんね。

大澤：うん。

斉藤：だから、IJCAI のペーパーとかが主なネタだったかもしれないですね。あとは本とか。最初の頃は私、イギリスで読んだのとか、ヴィノグラードのやつとかですかね。どうやって見つけてたんだろうね。

服部：シャンク[20]とかに見えますね。

斉藤：え？

服部：シャンク。

斉藤：シャンクね。シャンクも AI メモかな。何かちょっと分かんないな。スタンフォードと MIT と。アメリカだと、アメリカばっかりなんで、大体。イギリスもありますけど、アメリカだと大学が、どこの大学、AI やってるって、大体、分かってたから、そこの AI メモみたいなのを見て。その頃、まだインターネットはなくって。私、その最初の頃、インターネットの接続の仕事もしたので、それが 1986～1987 年かな、多分。

服部：そうですね。何か中島さんのときに、MIT 行ったときに、インターネット、まだなくて、ARPANET で本当に話を（　）。

斉藤：そうです。その ARPANET に最初、つないだんです。NTT の研究所も DEC の 20 っていって、まだあるかな。銘板が学生に見せる用に取ってある。とにかく DEC のマシンを買ったんですね、NTT の基礎研で。人工知能やるんなら、アメリカが使ってるのと同じマシンがいいよとかっていう話になって、それで DEC の 2020 って、最初、ちいちゃめの。ちいちゃめって、ちいちゃめってこれぐらいあるんですけど、その計算機、買って。その後、2060 って、もうちょっと、それが 3 ラックぐらい並んだような計算機、買って、私はそれの面倒もずっと見てたんですね。アメリカとつなぐ話も始まって、米国側に Mark Crispin[21]っていって、彼も死んじゃったんですけど、私より 2 つ下ぐらいかな。野島君と同じ年なんですけど、野島君も死んじゃったけど。野島久雄、ご存じですよね。Norman[22]の『誰のためのデザイン？』[23]を翻訳した。

大澤：そうそう。本で見たことある（笑）。

斉藤：そうか。彼、私、研究所にいたときの部下だったんです。厚木の基礎研で一

---

[19] Association for the Advancement of Artificial Intelligence (AAAI)、アメリカ人工知能学会。

[20] Roger Schank. アメリカの人工知能学者、認知科学者。

[21] Mark Crispin.　米国の計算機エンジニア。IMAP などメール関連のプロトコル開発など手がけた。

[22] Donald Arthur Norman. アメリカの認知科学者。

[23] D. Norman, 2002, *The design of everyday things*, New York: Basic Books.（＝2015, 岡本明、安村通晃、伊賀聡一郎、野島久雄訳『誰のためのデザイン？　増補・改訂版　―認知科学者のデザイン原論』新曜社）

緒に仕事してたんですけど。それで、何の話してたんだ？そう。DEC のマシンがあって、それでアメリカにつなぐ話をやってて、Mark Crispin とかって男が教えてくれて、何かちゃんとしたアドレスを取らないといけないんだと。グローバルアドレスっていうの取らないと向こうと通信はできないんだよって話を何か初めて知って、お願いして、もらったっていうのがやりだした初めで。その話は JPNIC のホームページの「歴史の一幕」っていう、何かシリーズがあって、その中に 2 年前ぐらいに書きましたから、そこに書いてあります。そうやってアメリカにつなぐのがスタートして、アメリカに行ってる人だとか。私、研究所に入ったとき、後藤滋樹[24]さんが直属上司で、奥乃さん[25]とかいろいろ、そこら辺にいたんですけど、後藤滋樹さんが音頭取りしながら、アメリカとつないで。最初、モデムですよね。電話かけてつないでたんですね、アメリカに。電話回線で。ある日、何か研究所の総務がすっ飛んできて「何でこんなに電話代、高いんですか」って言われて、後藤滋樹さんが一生懸命、説明してみたいな、そういう時代でしたけど。少しずつ、アメリカとのやりとりもできるようになって、奥乃さんもスタンフォードへ行ってたんだよな、何年か後で。向こうとやりとりしたりもしてたんですけど、その頃、ちょうどインターネット……。インターネットってまだ呼ばれてなかったけども、ARPANET なんだけど、インターネットとつなぐ話をやりだした頃で。

服部：いろいろ情報が得られやすくなった。その環境って何年ぐらいですかね。

斉藤：1987 年とか 8 年とかぐらいじゃないかな。だから、これをやりだしたときよりはもっと後なんですよ。

服部：10 年。

斉藤：もっと後なんです。これ、最初の頃は、そういう何か、あんまりメールが……。で、メーリングリストが出来上がったのは、ここにあるんですけど、実物が。実物ってのは変だけど、メーリングリストがあるんですけど。一番最初、見ると、メーリングリストの一番最初、見ると、86 年 5 月 19 日なんです。これ、AIUEO のメーリングリストのテストっていって、そのとき、10 人ですね。相田君[26]、伊庭君[27]、新田君、堀、堀先生ね、東大の。中川さん[28]、沼尾君[29]、小山。小山って誰かな。ちょっと待ってよ。小山ヒロコ、CS タイタンって書いてあるな。で、中島と、片

---

[24] 後藤滋樹。早稲田大学大学院理工学術院教授。情報工学。

[25] 奥乃博。早稲田大学大学院創造理工学研究科教授。人工知能。

[26] 相田仁。東京大学大学院工学研究科教授。高品質ネットワーク、並列・分散処理。

[27] 伊庭斉志。東京大学大学院情報理工学系研究科教授。

[28] 中川裕志。東京大学情報基盤センター教授。統計的機械学習、データマイニング。

[29] 沼尾正行。大阪大学大学院情報科学研究科教授。機械学習、データマイニング。

桐と、平賀君[30]がいたんですね。まだ彼、図書館情報大学かな。平賀譲ってご存じないかしら。その10人ぐらいで。まだ、だから、86年でまだ10人ぐらいなんですよ。もうちょっとたってから、わーっとなるはずです。要するに、まだ、まだ大学にいる人が中心。だけど、みんな、卒業してって、企業へ就職するじゃないですか、研究所でも。そうすると、そこの仲間、連れてきたりっていって、だんだん増えてくる感じですね、人数。

大澤：本当に最初の頃は、じゃ、修士というか、輪講会みたいな、学生……。

斉藤：学生中心でした。みんな、学生が……。ここの、だって、アドレス見たって、タイタンと、tokyo-uと、nttlab。だから、NTTには就職しだした人がいた、みたいな。私もそうですけどね。ぐらいのが一番最初で。メーリングリストは延々と、どこまで続くかっていうと、1998年まで続くんですね。その頃、松原君が、一番最後のメールは松原君が何か「僕が電子情報通信学会誌に書いたコンピューターチェスの解説が千葉大学工学部の電気電子工学科の入試問題に採用されたそうです。要約して、意見を書けという問題でした。事後承諾の形で、今日、実物の試験問題が送ってきました。うーむ、受験生がかわいそう」って書いてある。それが最後で。でも、この頃、ロボカップの話題が出てたり。そうそう、松原君がいつ AIUEO 辞めたかっていう話があって、中島は松原が結婚したときだって言うんだけど、だけど……。結婚したときは、ここに明確に書かれていて、91年だったかな。91年の正月か何かに彼がメールを送ってきて。ちょっと待ってよ。これか。違うな。A happy new year.　A happy new year. そうそう。「A happy new year」っていうメールを松原君が送ってきて、91年1月5日23時47分。みんな、夜更かしだよな。「明けましておめでとうございます」で、「ということで、私的なお知らせですが、4月に結婚することとなりました。ETL（電総研）ではこれまで隠していたこと、みんなに強く非難されたので、AIUEOの皆さんには取りあえず話を公開しておきます」って松原君が言ってきたのが91年なので。でも、その頃はまだ活発にやってたような気がしますけど、だんだんに何か、みんな、就職したり。私も実は、後でちょっと、ちょっとだけお聞かせできるかもしれない。私は最後だと思って、AIUEOで話した会があって、それが84年の7月のやつなんです。イズイットポシブルトゥーリプリゼントシンボライザーシンボリカリー[31]っていうやつで、その後は私、あれなんです。結婚したのもあるんですけど、フランス、行っちゃうんです、1年間。で、私はだんだんフェードアウト。その頃からはフェードア

[30] 平賀譲。筑波大学大学院図書館情報メディア研究科教授。認知科学、人工知能。

[31] 1984年7月21日のAIUEOでのプレゼンテーション「Is it possible to represent symbolizer symbolically?」のこと。

ウトしてきた感じで。でも、まだ続いてたと思うし、夏の合宿みたいなのは案外、その後もだいぶやってたような気がする。毎週のやつはもうあんまり開かれないんだけど、夏合宿ぐらいやろうよってみんな言って、あちこち。

服部：そういうふうに変わっていったっていうのは、最初は何か1週間とか2週間に1回。

斉藤：2週間に1回。これ見りゃ、分かるんですよ。これに日付が書いてあるから。2週間に1回だと思いますよ、多分、最初は。

服部：それがだんだん月1っていうふうに変わりつつ、合宿の、合宿でっていうような感じにだんだん移行していって。やっぱりそれは、あれですかね。最初は学生主体でっていうのが、だんだん就職していって、それで仕事をやりながらっていうふうになってくるので、やっぱり学生時代のような、密にやるのは難しい。

斉藤：そうですね。頻度は多分、そういうのが影響していて、中で話、して、どうしようかとかって言って、頻度……。だから、だんだん頻度が月1回ぐらいになってきて、みたいな。

服部：そうすると、それぞれ、立場が学生のときとは変わってきて、そうすると何か活動の内容自体もだんだん、何か興味の対象もよりばらけてくるとか、何かそういうので、だんだん活動が薄くなる。大体、普通、組織、大きくなってくると、だんだん活動内容が弱くなるというか、ばらけていってっていうこと、よくあると思うんですけれども、何か当初とはだいぶ、いつっていうよりは、徐々に仕事を抱える中でってなって、変わっていったようなところってのは、やっぱり感じるところはおありでしたでしょうか。興味を追求する、して、みんなでわいわいとやっていく雰囲気がだんだん弱くなっていくような。

斉藤：でも、若い人が入ってくるじゃないですか。若い人は学生さんだったりするんですよね。あとは社会人でも就職してすぐの新人とか。何か、だから、案外、最後のほうまで似たようなノリでやってたような気がします。年取ってくるから、われわれが。だから、われわれはちょっと、「そんなのは昔、あったよ」とか、何かそんな感じになるわけだけど、でも、若い人もいるので、若い人は自分が読んで「これ、面白かったんです」っていうのを発表してくれたり、「こういう研究したいんです」って話をしてくれたりすると、みんなでそれを面白いとか、つまんないとか、「こうしたほうがいいんじゃない？」とかって、そういう議論はしてたから、あんまりノリは最後まで変わってなかったんじゃないかな。

服部：そういうときに、その活動の中で、90年代、何かブームがまたやってきたりしてましたよね。そういう影響ってやっぱりあったんですか。

斉藤：ああ。いや、私は全然、全然でもないけど、ないと思っていて（笑）。っていうか、私はブームったって、やりだした頃はまだ第1次ブーム、終わってないぐらいかな。冬にもなってないん

だと思うんですね。イギリスでも、要するにそうやってアメリカの、さっきの勉強のあれみたいに、アメリカのやつを紹介したりして、AI、一生懸命やろうみたいな状態だったのを、日本に戻ってきて、やりだしたから、全然、冬ではなかった。それから、先生も、他の連中なんか読むと、先生が絶対駄目って言ったとか何か、松原も言うんだけど、私は全然、そんなことはなかったっていうか。私は修士は甘利先生なんですね。甘利俊一先生で、ほったらかしなので、基本。何にも言われなかった。卒論のときは和田英一先生なんだけど、和田英一先生はそんなに好きじゃなかったかもしれないね、AIは。

服部：そういうふうに、中島さんはおっしゃってました。

斉藤：うん。だけど、和田英一先生のときは、私はAIじゃなくて、BCPLっていう言語の処理系つくったので、全然、AIじゃなかった。知らなかったです。学部のところまで、AIっていうもの。エセックス行って、知ったので。　だから、外の影響っていうか、周りが何やってるかとかっていうのは、そんなに振り回されなかった。ICOTに随分、人が行っちゃったからね、確かにあのときは。10年ぐらい続いたでしょ、ICOTって。

服部：そうですね。

斉藤：そうすると、みんな、だいぶ行ったけど、私は全然、興味がなかったっていうか、何か、何か違うなって最初から。趣意書みたいの。知識工学自体も何か違うなって思ってた。

服部：ちょっと話が戻るようなんですけど、今の（　）ありましたが、やっぱりそうすると、もともと計数の学科ですよね。計算機は興味があったけれども、AIなんていうのは、そもそもその言葉自体がそんなに認知されてないような。本当に何か、例えば偶然みたいなものですか。

斉藤：まあ、そうでしょうね。分かんないな。その前に知ってたか、人工知能。やっぱり知らなかったと思います。全く真っさらで、何も知らなかったけど、エセックス行って、授業を選ぶときにそういう項目があって、解説、何か説明が書いてあるわけじゃないですか。人間のインテリジェンスを計算機使って、解明していくって。わあ、面白そうとかって思って、取って、授業も面白かったし、さっき言ったみたいに、輪講みたいので、案外、ディスカッションになると難しいんです、ヒアリングが。だけど、だんだん。最初、大変だったけど、だんだん聞こえるっていうか、分かるようになってきて、そうすると、面白い議論をいっぱいしてたので。exampleとか、面白いんですよね。だから、赤ん坊の。1つだけ覚えてるのは、それ、パット・ヘイズが言った話。マイケル・ブレディーが言ったのかもしれないけど。どっちか覚えてないですけど、ビジョンの話で、目から……。要するに立方体を見せると、人間は全部、辺の長さ、同じだと思うわけよね。だけど、ちっちゃな子どもにこう近くで見せると、こっちのほう

がこっちより長いって言うっていうわけ。えっ、本当かなと思って、いろんな子どもで実験したんだけど、本当にそのとおりで。つまり、遠近のあれが出来上がるのは、っていうか、立方体という知識が入ってから、見た、見た画像は、プライマルスケッチみたいなやつは、ちゃんとこっちが長くて、こっちが短くできてんだけど、それ、同じ長さだって言うようになっちゃうのが後知恵なんだよね。とか、そういう話をするわけね。何かビジョンの議論とかのなかとかで（　）。そういうのがとっても面白くて、一生懸命、付いて、付いてって。

服部：そういった話をいろいろ聞いた。で、帰国をされますよね。もう本当に何か、誰っていうわけじゃなくって、とにかく面白いから、何か一緒に勉強しようって。

斉藤：みんなにアナウンスしました。情報輪講って何人ぐらいいるのかな。大きな部屋だったから、20 人ぐらいいるのかな。学生も。学生だって。ちょっと待てよ。修士、ドクター、全部いるのかな。情報輪講ってそうだったかな。先生もいて、みたいな。

服部：もう、あれですね。そうすると、何かはやりだからとか。

斉藤：全然違う（　）。

服部：そんなんではなくて？

斉藤：だって、手、挙げたの、3 人しかいないわけだから。それって、はやってないってことだよね、多分。

服部：でも、何か面白そうって、本当に興味で、ふぁっと集まったのが、最初の 4 人？

斉藤：そうですね。

江間：手、挙げた、中島先生とか片桐先生っていうのも、もともと全然関係ない研究されてたってことなんですか。

斉藤：いや、でも、AI といっては、やってなかったかもしれないな。中島なんか、処理系のほうの話、してただろうし、片桐は何してたんだろう。でも、言語だったような気がする。言語やってても、AI じゃない言語、いくらでもあるじゃないですか。そういうこと。白井君もそうだよな、確か。だから、その場でみんな、何か面白いと思ったか何か。少し知ってたのかもしれないし、AI の、アメリカでは少しそういうのやりだしてるとかっていうの、知ってたのかもしれないし。

江間：何か、でも、今、始まりのお話で、先ほどもずっと長く続いていく中で、今、新陳代謝っていうか、人の入れ替わりも出てくるっていうのと同時に、やっぱりいろいろ輪講で読んでいると、AI の定義とか、技術とか、内容、どんどん変わっていくわけなんですかね。誰が何、取り上げるかにもよるかもしれない。先生は途中でフェードアウトされたっていうふうにおっしゃったんですけど、他にフェードアウトされる方とか、入れ替わりは結構、激しかったんですか。「何かこれ、違うな」みたいなふうな話とかも中でされたりとか。

斉藤：そういう議論はいっぱいしてましたけど。

江間：何が AI かみたいなのは、みんな、そ

れぞれ違ったんですかね。それとも、何か共有できるものはあったんですかね。

斉藤：私はこないだ、FUN-AI 行ったときに、FUN-AI 行ったときに感じたのは、最近は何か随分、変化してるような。変化っていうか。だから、第 3 次ブーム、ほとんどニューラルネットじゃないですか。それって、AI とは全然、全然ではないけど、だいぶ違うと私は感じていて、ICOT もできて、知識工学だとかって言ってたときでも、AI 自体の定義っていうか、もともとの目的とか、何を研究するんだっていうの、あんまり変わってなかったような気がして。それは私の場合は本当、何かダートマス会議[32]なんですよ。ここにあるんだけど、この本、この会議のときに、みんなが思ってた。これ、マッカーシー。すごい人、いっぱいいるんだよね。マッカーシー[33]とか、ミンスキー[34]とか、シャノン[35]がいて。

江間：シャノンとか。サイモンもいる。

斉藤：え？

江間：サイモン。

大澤：ハーバート・サイモン。

斉藤：サイモン？　サイモン、サイモンね。サイモンは何か、そうそう、出るんだけど、途中で帰っちゃったみたい。ちょっと何か学派、違うふうの。でも、このときに、これ、マッカーシーが言いだしたんだよね、言いだしっぺは。それで、マッカーシーはやっぱりっていうか、その後、みんな、シャノンもそうだけど、マッカーシーも若干、そういうとこあるけど、チェスの研究、始めるわけよ、計算機。つまり、まだ計算機はすごく遅いし、全然、まだそういう、今のような状態ではないんだけど、でも、原理的に知能って、計算機でできるぞと、実現できるぞって思った人たちが集まって、やろうよって言いだしたので、私はそのノリとほぼ一致していて、今でも信じてるけど、絶対、計算機上に実現できると思って。時間はやたらかかるだろうけど、人間の知能と同じようなものが……。で、やり方もまだ分かってないと思ってるけど、いつの日か実現できると。それを目指して、何かいろいろ工夫してるっていう意味では何も変わってないんですけど。

江間：それを結構、じゃ、AIUEO に参加してた人の中では、あるいはコアメンバーの最初の……。

斉藤：最初の頃のメンバーはみんな、共有し

---

[32] 人工知能に関する会議 “The Dartmouth Summer Research Project on Artificial Intelligence”の通称。1956 年の 7 月から 8 月にかけて開催され、「人工知能（artificial intelligence）」という言葉がはじめて使われたとされる。

[33] John McCarthy. アメリカ合衆国の計算機科学者。人工知能研究の第一人者で、プログラミング言語 LISP の開発者としても知られる。

[34] Marvin Minsky. アメリカ合衆国の計算機科学者。ダートマス会議の発起人の一人で、「人工知能の父」とも呼ばれる。

[35] Claude Elwood Shannon. アメリカ合衆国の電気工学者、数学者。情報理論の父として知られる。

てると思うな。だから、最近の話を聞いても、みんな、別にね（笑）。

全員：(笑)。

斉藤：いやいや、そうねと。フィーチャーディテクターがうまくできるようになったのね、みたいな。

江間：そのオートマタスタディーズ[36]は読まれてるんですかね。

斉藤：この中では読んでないね。これ、私、どっかで、古本屋で見つけたんだよ、神田の。

大澤：神田にそんなものが。そうですよね。神保町、すごいな。

斉藤：本当に。神保町、すごいよ。

駒谷：50 年代ですか、出たのは。

斉藤：え？　これ、出たのはね。

駒谷：56 年じゃないですか。

斉藤：出たのは 1957 年。あ、セカンドプリント。56 年。56 年。6 ドルだもん。

大澤：初版。2 版、じゃあ、2 版、すごい。

斉藤：初版なんだか分かんない。でも、これ、買ってきた人がいたとか、そういうことだよね。古本屋。

大澤：そうなりますね。

江間：しかも、売っちゃったのね（笑）。

斉藤：それで、面白いのは、誰のだっけ。もう 1 人の、橋田じゃねえや。インタビューしたの。

江間：堀先生。

斉藤：堀先生の読んでたら、インテリジェンスアンプリファイアの話が出てきたけど、インテリジェンスアンプリファイアって、ここに出てくるんですよ。

大澤：なるほど。

斉藤：初出はここ。アシュビー[37]だか何かが書いてて、それは本当に。っていうか、もうちょっと広い意味で、人間と同じようなことが計算機にできるか、できないかをちょっと理論的に解析してるような感じのペーパーなんだけど、「Design for an intelligence amplifier」[38]っていう、アシュビーが書いてて。それ、そういうのを見てたから、だから、私は AI は、っていうか、人間の知能は絶対に計算機で実現できると思っていて、時間はすごくかかると思う。今でも思ってるけど、そんな簡単にできないけど、それを目指して、頑張ってくのが AI だろうと。

服部：やっぱり特にコアメンバーが、戻るところってやっぱりそういうところで、それ、共有できる人は大体、コアになっている。多少、何かずれはいろいろあると思いますけど。その中で、例えば何かあるところ、かなり哲学なんか、とこ、絡んでいたところもあるけれども、それも人間の知能とか知性っていうのは何なのかっていうのを理解する上で、そういう議論っていうのは、すごく活発になってできていたっていうところはある。

斉藤：でも、あんまり哲学議論をした記憶はないんだよな。いや、哲学議論って変だけど。つまり哲学の人と話すると、

---

[36] C. E. Shannon, J. McCarthy, 1956, *Automata Studies (Annals of Mathematics Studies)*, Princeton University Press.

[37] W. Ross Ashby. イギリスの医学者。サイバネティクス、複雑系の研究で知られる。

[38] W. Ashby, “Design for an intelligence amplifier.” Automata studies 400 (1956): 215-233.

もう何か言葉の定義みたいなのになったり。土屋俊なんか、あるとき「うそはつくためにあるんです」って言いだすから「何、それ」とかって言いだして、みんなでやり込めたんだけど。でも「うそはつくためにあるんだよ。分かんないのかい？」って、とか言うから（笑）。何だろうとか思って。そういう議論はあんまりしなかっ……。いや、みんな、やっぱり理系だし、計算機はみんな、好きなメンバーだから、計算機でつくるのにはどうしたらいいか。そこの難しさは何かとかっていう議論がメインで、そんなに。でも、読んだのの中には哲学みたいなのも読んだけどね。

服部：何かそれぞれの立場で、哲学者から見たときの人工知能、計算機をやってる人間から見たときの人工知能。真ん中に人工知能っていうのがあって、興味を持つ観点って、多分、違うんだけれども、ただ、人工知能ってどういうものとか、どう考えればいいのっていうのを考えるところは、違う観点から、いろいろがくがくとやるっていうところは、すごく活発というか、楽しめたというところですかね。

斉藤：うん。そうね。

服部：何か、計算機のほうから来たときに、哲学っていうのは、何か堀先生のときに、中島先生のときにも聞いたかもしれないんですけど、哲学をやってる人と、そんなに簡単に話ができるものかなっていうのをやっぱり。この集まりの最初のときに、何か言い方が何か荒っぽいみたいなところをとがめられたりしたことがあったので、そんな、そんな深い議論をそんなやっていけるかっていうのは、結構、やっぱり気になるんですよね。怒られました、最初。

斉藤：私も何かでね。あれは何だっけな。何かで誰かにこっぴどく怒られた記憶があるんだけど、何で怒られたか、全然、もう覚えてない。それは堀君のに出てきた、ほら、何だっけ。寺子屋っていうやつかなと思ってたら、そうじゃなくて、NTT は NTT で情報文化フォーラムとかっていうのを、2000 年ごろだけど、やっていて、何か本も。松岡正剛[39]なんです。松岡正剛に NTT が何か依頼して、そういう、ちょっと松岡正剛流、何でもありみたいな議論をする場があって、その中で AI のサブグループみたいのがあって、研究所から呼ばれて、私とか、竹内郁雄さん[40]とか、後藤滋樹さんも時々出たかな。中島も時々、1 度か 2 度来たんだけど、彼はやっぱり哲学……。哲学議論はそんなに好きじゃないんだと思う、中島は（笑）。それで、そういう文系のね。文系っていうか、本当の哲学の人たちの何か訳の分かんない議論聞いて、「つまんねえ。俺、やめた」みたいな感じ。私は何か、ああいう人がああいう議論をすること自

[39] 松岡正剛。編集工学研究所主宰。

[40] 竹内郁雄。日本電信電話公社電気通信研究所、NTT 基礎研究所、NTT ソフトウェア研究所、電気通信大学、東京大学を経て早稲田大学理工学学術院。

体に興味があり。いや、つまり、それで何のアウトプットを出そうと思ってるのかなとか思うよね。いや、ごめんなさい（笑）。文系の人にはあれかもしれないけど、何が目的なんだろうとかって思わなくもないけど、でも、根本的な何が真実とか、人生の目的はとか、そういうのも議論してもいいけど。でも、AI とはその頃はそんなに切りむすばなかった（　）。大体、途中で第 2 の冬の時代かな。要するに AI 批判がいっぱい出てきたときに、批判する人たちは哲学者だったから、批判されたら、何か反論するためには、哲学をちょっとはかじっとかないと反論もできないよな、みたいな感じで、ちょっとは勉強したみたいな感じ。

江間：批判した哲学者って、どういう人たちですか。

斉藤：どういう人とか、みんな、そうだよね。

江間：国内でじゃなくって？

斉藤：国内にはあんまりいなかった。黒崎さん[41]なんか逆に、どっちかっていうと「AI、面白いね」っていう感じだったから。批判した人、土屋俊はどっちなんだろうな。

服部：（　）するのが好きですよね。

斉藤：よく分かんないけど、でも、あんまり批判的じゃなかったと思う。批判してる人は……。批判すらしない。「何だ、あれは」って思ってた人が多かったのかな、日本の哲学者は。

大澤：あんまり注目されなかったってわけではないですか。哲学者の人にそもそも、人工知能。

斉藤：人工知能が？　ああ、日本だとそうかもしれないね。だって、成果がそんなに出てるわけじゃないからね。

大澤：何か黒崎さんも、私も読んだ記憶ありますけど。

斉藤：何、『電気羊はアンドロイドの夢を見るか』[42]っていうやつ。

大澤：確かに、そうですね。高校ぐらいだった記憶があるんで、多分、その辺だと思うんですが（笑）。あんまり何かそれ以外で、人工知能ってのはそもそもワードに出てこなかった記憶がある。

斉藤：そうかもね。

神崎：2000 年ぐらいだと、情報倫理って、インターネット社会とかに対する批判とかはあって、その一部として情報、人工知能の話を扱った人もいたでしょうけど、そんなにメインの話題になったっていう印象はないんだろう。

服部：2000 年は第 2 次も一段落ついてっていう頃だから、もう AI、むしろ AI って言うなっていうような（笑）。

斉藤：そんな雰囲気があったの？

服部：何か、人工知能のモダンアプローチの本[43]、あれを出すときに……。ああ、そうですね。あれです。あれを、何か訳本を出すときに、AI という言葉を使わない。AI って言うと売れないからって言って、で、何かエージェント

---

[41] 黒崎政男。東京女子大学現代教養学部教授。哲学。

[42] 黒崎政男の 1987 年の著書『哲学者はアンドロイドの夢を見たか　人工知能の哲学』（哲学書房）のこと。

[43] 1997 年出版（第一版）のラッセル他、古川 康一（監訳）『エージェントアプローチ　人工知能』（共立出版）のこと

って付けたとか。何かどっかで聞いた覚えがある。AI だけだと駄目。何かそんなような話を聞いた。

斉藤：そうなんだ。

服部：エージェントアプローチってどこから来たのかな、みたいな。そしたら、何かどうもあえて付けたという。

斉藤：だから、みんな……。だから、私は多分、ちゃんと自分で AI をやれてないんですよ、実は。つまり、エセックスではかぶれ、帰ってきて AIUEO もやってたけど、自分でメジャーな成果は何も出してないので、AI に関しては。そうなんです。それで、私はもう、最後の AIUEO のときに、こうやってプログラム書いてたの、しょうもねえじゃんっていう話をして、どちらかっていうと、認知科学のほうにシフトしてっちゃうんですね。だから……。でも、認知科学も全然、駄目だったわけで、（ ）だけど。

服部：AI から認知学にずれるときっていうのは、そのときの AI っていわれている、これ、AI でやるんなら、これだよねって言われてるものが、元のこういうやりたいことから乖離していったものになってるとか。

斉藤：いや、そうじゃなくて、やりたいと思ってる。だから、人間はどうやって考えてるんだとかっていうことをアタックする手段として、プログラムを書いてみて。モデルをつくったり、それをプログラムに書いてみたりってやってることが成果に結び付かない。ってか、頭ん中が分かるようにならないっていうジレンマっていうか。デビッド・マーも、ちょっと違うんだけど、そういうような批判をしてたり、ウィノグラードも。私、ウィノグラードとか、デビッド・マーとか、誰が好きって言われるとそういう人たちが好きだったんだけど、ウィノグラードもあんなすごいプログラム書いたのに、やめて、離れてくよね。それは理由が多分あったんだと思うし、私が思うに、それはやっぱり、そうやってプログラム書いても、あれはトイプログラム[44]だし、すごいトイプログラムだと思うけど、トイプログラムだし、実際に人間がそうやってるのかって分かんないっていうのが背後にあったんじゃないかなと思って。

大澤：そこはやっぱり、当時ってか、その辺の人たちは、やっぱり人間をモデルにして、それを知りたいっていうのは結構強かったということですか。

斉藤：うん、と思うけど。今でもそうじゃない？

大澤：いえ、もちろん。もちろんそうなんですが。でも、今、結構いろんな。つまり、知能そのものを扱いたいって人も。そうですね。

斉藤：知能そのもの？

大澤：インテリジェントなシステムっていうもの自体。

斉藤：人間が脳の中でやってることではな

[44] 複雑な現実世界ではなく、単純化され限定的なルールがある世界での問題（＝トイプロブレム）を解決するプログラムのこと。

く。

大澤：人間もやってるかもしれないけど、人間以外もやってるかもしれないみたいな、そういうゼネラルな、多分、話。

斉藤：ゼネラルインテリジェンスって、でも、サイモン、ニューウェルを批判する人は、サイモン、ニューウェル、ゼネラルプログラムソルバーだから、あれがゼネラルな枠組みだって。つまり彼らは本当に人間だけではなく、プロブレムを解くっていうのはこういうことだって言って、Means-Ends Analysis とか何だとかって言ったわけで、それを、しかも一応、実験をして、人間で。ほんで、実証したわけよね。実証……。ゼネラルなセオリーっていう意味では、そういうの、もう大昔からあって。でも、それで何もできなかったっていうか、例えば私はエセックスの修士は暗算の天才と同じことをやるプログラムを書くっていう課題だったんだけど、プロダクションシステムを使ったんだけど、もう全然、違うよなって感じたんです、やり終わった後で。だって、要するに頭ん中で本当に起こってることと、そのプログラムが動いてやってることと、表面的には一致するんだけど、それって、本質的なところで strong equivalence がないとかって言ってたけど、そういう感じがして。だから、プロダクションシステムみたいな手法とか、プログラミング言語とか、そういうものをいくら駆使しても……。だから、私は何か脳の研究とか、認知科学の研究とか、そっちでもうちょっと中のことが分からない限り、できないと。いや、本当にインテリジェントなシステムはつくれないというふうにすごく感じていて。でも、最近になって、囲碁。私は、それで囲碁の研究、少ししだすんだけど、あんな強いプログラムが、ああいう手法でできてしまったので。

大澤：そうですね。brutal なというか（笑）。

斉藤：brutal だよね、考えたらね。

大澤：ええ？っていう感じ（笑）。

斉藤：brutal 以外の何者でもないよね。ほんで、ちょっとあの日はショックだったけど、何か話を聞いた日は、1 日。老後の楽しみを奪われたってね。

大澤：あまりにも急でしたよね。何かちょっと（　）。

斉藤：いや、そうね。スピードは予想外だったけど。でも、でも、みんな、プロ棋士の人たちもそうだけど、やっぱりあれじゃ分かんないって言うわけよ。つまり、説明。人間同士がやってると、でも、人間のリポートがどれぐらい正しいかも、分かんないけど、一応、頭の中でこう考えたから、ここ、打ったんだとかってか、目的とかそういうレベルの概念が必ず入ってるわけよね、ゲームするときでも。目的なしにゲームはしないから。しかも、いっぱい、サブ目的があって、それがインタラクションしてみたみたいなことが起こってるわけ、頭の中で明らかに。そういうの全くなく、関係なく、単に画像として見て、「こっちが良かったんです、はい、それでおしまい」になっちゃったから、あれ、つまんないよね。

大澤：そこはやっぱり、人間がやってるプロ

セスが分かんないっていうことがやっぱつまんないってこと（　）。

斉藤：それを知りたいんだもん、だって。それが最終的な目的だと思ってるから、今でも人工知能の。

服部：やっぱりあれですよね。ディープ……。チェスのときと結局、同じですよね。

江間：ディープブルー。

斉藤：ディープブルー？

服部：あれのときも、何かチェスのグランドマスターがどういうふうにチェスをプレーしてるのかってのが分かったわけじゃなくて、あれも結局、計算機の力で勝ちましたっていうところが分かったときに、僕も学生だったので、最初は何かそのニュースしか知らないわけですよね。何かこういう写真だけ見て（笑）。

江間：カスパロフが？

服部：負けたとか。そう。負けたかって思うんだけれども、でも、何かいろいろ見てると、別に何かそういう、人間、人間がどう打ってるのかって分かったわけではなくって、とにかくたくさん、計算機、走らせましたと。勝ったと。読みの深さで勝ちましたって、確かに。僕、そのとき、学生だったので、何かちょっとそういうのじゃなくって、どう。賢く打つって、どういうのかっていうのは分かんないのかって思った感じはあります。あれと同じ。印象としてはやっぱり同じなんですよね。確かにね。

江間：何か最近では結構、DARPA とかが説明可能 AI、XAI とかいうふうに言いだして、お金付けて、韓国も始めたりとかしているじゃないですか。あそこで目指してるものと、先生の言う、中を説明してくれるのはまたちょっと違うんですか。

斉藤：それは違うね、全然ね。だって、今、言ってるのは、ニューラルネットで仕組みをつくっちゃうと、ニューラルネットの中のパラメータしか分かってなくて、データとしては。何にもないから、それだと出した答えをちゃんと説明しないと、安心、人間が安心して使わないよっていう、そういう議論でしょう？　だよね、多分ね。そうでもない？

大澤：人間に理解可能な形で説明（　）。

江間：可能な形で。

斉藤：提示しなくちゃいけないし。

江間：だから、何か透明性みたいに、中のアルゴリズム、全部出せ、みたいのとはまたちょっと違う方向ですけど、でも、おっしゃるとおり、何か内部処理がどうとかいうよりは、何か人間はどういうふうにしたら理解できるのかなっていうのを考えた上で、こう出したら、納得してくれるでしょ、みたいなのも入ってくると思って。

大澤：解釈機っていうか。

江間：解釈機のリーズンを出すっていう。それは本当に中で起きてるのかどうかはまた別の話かもしれないけど。

大澤：批判はありますよね、それが要するにだましてるんじゃないか。人間にこう、こうやってるように思わせれば、感じっていうのが、本当の説明と 1 対 1 対応してるのかっていうのは来るので。

斉藤：そうだよね。でも……。

大澤：ただ、リーダブルなものをやる商業的な動機にはなってるかなとは思いますけど、なので。

斉藤：うん。だから、全然違うと思います（　）。

大澤：全然違うのかな、確かに。

斉藤：いや、だから、何か脳科学っていう。何か変な脳科学の人がマスコミにはいっぱい出てきてあれだけど、そうじゃない意味で、本当に脳がどう、何を処理してるかを知りたいっていうの 1 手段として計算機を使うっていうのが有効なはずだと思ってやりだしたけど、あんまり有効ではなかった、みたいな。まだ全然、分かんないですよ。っていうか、ニューラルネットでうまくできるような下のレベルの処理と、人間も言語化して、言葉で考えてるって、みんな、よく言うよね。本当にそうかどうか分かんないけど、言葉で考えてるレベルの処理とあって、言葉で考えてるレベルの処理は、別に今までのプログラムとかプロダクションシステムで書いたのとほとんど同じかもしれない。頭ん中で推論してるときとかは。分かんないよ、違うかもしれない。だから、今までの、最初の頃のそういう AI が無駄だとは私は思ってなくて、ニューラルネット出てきて、っていうか、出てきたのはすごい昔だけど、あんまりうまくいかなかったのがうまくなった結果、私が昔言ってた、シンボライザーって言ってた、外からの入力から中で処理するシンボルになるまでのところが、もしかしたらある程度、つくれるようになってきたかもしれないって気がしていて、それと、そうやってつくったシンボルと上とをちゃんとつないで、AI のシステムをつくったら、それこそさっきの意味で、本当の意味で説明ができる。相手から質問されたら、シンボルの基のほうまで下がって、「いや、でも、見てごらんよ。ここのこれがこうだから」みたいなことが言える AI ができるかもしれないじゃない。そういう気はすごくしてて、可能性、すごく感じてるんだけど、もう年取っちゃったから、駄目かもしれないよ。

久木田：ちょっと今のシンボルの上っていうのが何なのかが、ちょっとお伺いしたいんですけど、シンボルっていうのがあって、それを上から動かしてるのが、何かプロセッサーみたいなものですか。シンボルプロセッシングでも（　）。

斉藤：例えば LISP でもいいけど、LISP という処理系は人工知能のためにつくられたんだけど、シンボルの塊だよね。っていうか、シンボルをマニピュレートして（　）。頭ん中で同じことをしてるかどうかは別として、人間も、だって、言葉を使って考えてるっていうことは、言葉は頭の中でいろいろぐるぐる回してるよね。処理、処理してる。言葉にならないシンボルも私はあると思うんだけど、そういうの、うまく組み合わせて、何か頭ん中でやってるレベル。言語化できる処理はほとんど計算機の昔のそういう LISP でシミュレートするっていう、プロダクション

システムでプログラム書いて、同じようにやるっていうのと、あんまり変わらないと思う。だって、お医者さんだって、知識はだいぶ言語化された知識で、何か病気を診断するときに幾つかチェック項目があって、これがイエスか、これがノーかとかってやって。だから、じゃ、こいつは何とか病だなとかってやるっていうのは、もうMYCIN[45]がやってたのと全く同じことで、でも、人間の頭の中でそれ、やってるよね。患者の話、聞きながら、多分。そういうのはプログラム書いて、それと同じようにやって、全然いいじゃん？　いいっていうか、それ、人間のやってることと同じこと、計算機でやってるんだと思うけど、そこで使う、そのシンボルみたいなものをビジョンとか、耳から入ってきた音の情報とか、触った感覚とか、そういうのを全部、統合して、そういう中で扱えるシンボルにするところまでが大変だっていう話をずっと言ってたんです、私、昔。そこ、シンボライザーと呼んで、そこが難しいんだって。あるいはそこのほうが知能全体で考えると大きな割合を占めていて、そこをちゃんとつくらないと、ロボットでも何でも人間のように柔軟に、外界にうまく適応して動く知能がつくれないとかって思ってて、今でも思ってますけど。

久木田：松尾先生[46]とか松原先生はその辺が今の深層学習でどうにかなるって感じの意見じゃないですか。

斉藤：まあ、そうかもしれない。

大澤：そうですね。

＜電話＞

斉藤：すいません。失礼いたしました。

大澤：何か、いわゆる例えば今だと医療系で、それこそディープラーニングみたい、がんの画像出して、よく判定できるみたい、ありますけど、ああいうのも判定してるって意味ではシンボル化はしてますけど、そういうレベルじゃない話ということですかね、そういう意味では。

斉藤：画像見て、がんの判定……。え、何？シンボル化の定義？

大澤：定義というか、例えばそういうレイヤーではできてるかもしれないけど、そこではない。でも、それ、人間がそれやってるのとは違うという感じですか。

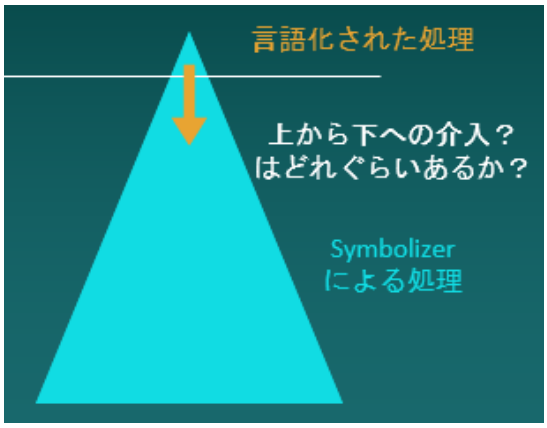
斉藤：そのレイヤーは、案外、ローレベルだと私は。つまり画像を見て、これ、怪しいぞとかっていうのは、それ以上、説明できないじゃない、人間は言語で。そういうのって、だから、ビジョンから始まって、シンボライザー、シンボル化されるまでのところでできちゃってること。人間、何ていうのかな。シンボル、こういう絵をよく描いたんだけど、ここから入力が来るわけね。音でも、ビジョンでも何でもいいんだけど。中でこう、いろいろ統合されてる。ごちゃごちゃして、最後、ここら

---

[45] 1970年代にスタンフォード大学で開発されたエキスパートシステム。

[46] 松尾豊。東京大学大学院工学系研究科技術経営戦略学専攻消費インテリジェンス寄付講座共同代表・特任准教授、産業技術総合研究所人工知能研究センター企画チーム長。人工知能。

辺でやっと言語化できる。ここで何かやってる言葉が言語化できるんだよ。



図：シンボルとシンボライザー

大澤：なるほど。

斉藤：だけど、この氷山の下のように、ここにいっぱいあるものは言語化できない。だって、頭ん中でアクセスできない。何を言語化してるかって、とっても不思議だと思わない？　人間。聞くとちゃんと答えるけど、私、囲碁のその実験やってたときのプレーヤーに「何を考えてるんですか」って聞くと、みんな、ちゃんと答えるんだけど、それって、ここら辺だけでぐじゅぐじゅ言ってて、ぱっと見たときに……。だって、それで実験やると、詰め碁ってのがあって、詰め碁を解くときのアイカメラ付けて実験やると、アイカメラで、こうしゃべってくれる前にぱっとその答えのとこに目が行ってるんだよ。そのときの処理って、ここ何にもやってなくて、ここの下で何かできちゃってるんだよ。

江間：無意識（　）。

斉藤：こういう部分の一部がニューラルディープラーニングとか何とかでできるようになったっていうのが私のイメージで、だから、逆に言うと、これと上の言語での処理。昔の古典的なAI とうまく結び付けたら、もっと面白い AI ができるんじゃないのっていう気がしてる。

大澤：私は、そういう意味でインタラクションをやっている人間なので、言語化できるってのは、つまり他人と共有できるって意味なのかなと思っていて。

斉藤：まあ、結果として。

大澤：そうですね、共有。例えば何でしょう。すごい、要するに言語化できないとこを無理やり言う人もいて、何か白い大きな光があって、それがこう、どうこうと、将棋の棋士でもそういうこと言う人いますけど、それはつまり共有できないあれに基づいて言ってるから、何もないけど、シグナルになるところっていうのは多分、ある程度、共有化されてる。

斉藤：でも、ここ、ここ、何ていうのかな。人によって、どこまで言語化できるかは、何か随分、違うような気が私はする。例えばイチローって、すごい何か筋肉アウェアな人だよね。説明とか聞いてると、自分の筋肉に関して、すごくよく分かっていて、今日はここの何とか筋が……。いや、普通の人は全然知らないような、体の、自分の体の内部のことを、ここから挙がってくる情報なんだと思うけど、何か把握してて、じゃ、今日はそこを鍛えてやろうとか、そこが疲れてるから、そこみたいなことをよく言うじゃん。あれ、そうやって訓練……。囲碁のプレーヤーもそう

なんだけど、そうやって訓練すると、ここの下の、これ、すごいと思うんだけど、膨大な、膨大っていうか、処理としては短時間で終わっちゃうから、不思議なんだけど、でも、ここの仕組みに関して、上からアクセスできる。

大澤：ある程度、アクセス。

斉藤：ある程度、アクセスできる人もいるような気がするんだよ。その人のプロトコルとかをうまく取ると、ここの研究が少しできるかもしれないなと楽観的に思ったりする。

大澤：そこは割と、でも、本当に認知科学。

斉藤：まあ、そうだね。

大澤：面白い。なるほど。

斉藤：そうしないと説明もできないっていうか。だから、囲碁の人たちだって「勝手にもう直感です」とか「第一感、これなんです」とかって言うわけじゃん。その第一感はどこから出てきたのって言ったら、大昔から何か、たくさんプレーしてて、その間にためた、それこそビジュアルイメージみたいなものとか、形の良さ。形がいいとか悪いとか、すごく言うんだけど。

大澤：ああ、ありますね。

斉藤：形がいいパターンだよね。そのパターンがもう頭に染みついてるわけだ。下のほうの処理。そういうのなしには上のことはできないし。だから、逆に言うと、ディープラーニングじゃない、AlphaGo で衝撃的だったのは、ほとんどそういうレベルの処理だけで、囲碁は強くなれちゃうっていうゲームでしたってことだよね。

大澤：タスクの性質が分かった（　）。

斉藤：いや、それ、一生懸命、人間は何か言葉をつくり、お互いに議論し、ああだこうだってやってたんだけど、そんなのほとんどなしで、その誰よりも強くなれるようなゲームだったってこと。

服部：（　）把握して、こういうふうに言ったら、こうなるのでとかじゃなくて。

斉藤：でも、人間はもう思考能力が限られてるから、だから、そうやって問題を分割し、ディバイドアンドコンカーでやるしかないじゃん。だから、人間の対局見てると、あれなんですよ。あるところが一段落するまではそこをやるんだよね。順序はどっちでもいいっていうのもあるけど、でも、AlphaGo の打ち方は「何でここへ打つんですか」って、みんな、言うわけだけど、プロの棋士の人たちは。関係なく、部分も認識してはいないし、その優先順位も考えてはいないし、最終的な答えが、自分が勝つか勝たないかだけを。AlphaGo の仕組みは少し知ってるかな。そういうやり方で結論を出してるわけだから、平らになるはずないよね。

大澤：何か書かれてましたけど、モンテカルロの方法自体が既に結構、画期的というか、（　）。

斉藤：いや、私はあれ、画期的だと思うけど。あれ、言いだした人は物理学者なんだよね。

大澤：ああ、そうか。

斉藤：それで、全然、囲碁を知らない人が最初に。言いだしっぺは、だよ。後で精緻にしてったのはみんな、研究者なんだけど、最初に言いだしたのは物理学者で、何か Simulated Annealing って

いうの。要するに物質系がエネルギー最小の状態に落ち着くよね、大体の物質は。その最小の状態を見つけるための手法があるんだけど、それと同じようにして囲碁ができないかって思ったんだよね。すごい、すごいな（　）。

大澤：あれがすごいなと。いや、つまり確率的に打ってけばいいじゃないっていう、結構、乱暴な方法ですね。

斉藤：乱暴だけど、それをたくさんやるから、つまり統計的に意義が出てきて。

大澤：あと、囲碁ってゲームがそれに合ってたのかなとは思うんですけど。石の打ち方。

斉藤：そうだね。画像っていう意味でもそうだよね。本当に最近のニューラルネットの進歩にぴったりマッチしてる分野で、そこにちゃんと目をつけてやったのがやっぱり偉いね、DeepMind 社。

大澤：何か言語化する方向とは違うような気が何となくするんですよね。何だろう。

斉藤：そうだよね。言葉って、だから、シリアルにしなくちゃいけないし。

大澤：うん。そこは今、汎用（はんよう）人工知能研にいるんです。よく議論するんですけど、私は結構、ストロングに、言葉っていうのは他人と共有するために使ってるものだから、そのレイヤーでできることは結構、限界があるんじゃないですかっていうふうに言うんですけど、別に他人とコミュニケーションしなくても、ロジカルに頭の中で組み立ててる面もあるんじゃないのって話もあって、よく議論はするんですね、そこで。どの辺までグラウンディングできるのかでちょっと興味があります。人間のやってることっていうのがシンボルのレイヤーで。

斉藤：でも、中でもだいぶ使うよね、やっぱりね。

大澤：うん。とは思うんですけどね。でも、2 次的な目的かなと思ったりするんですけど。

斉藤：2 次的な目的？

大澤：そうです。つまり進化的には、別にそれが合理的なわけではないかなというような感じですかね。何でしょう。

江間：あと、何か言語って言っても、日本語と英語で違いあるのかとか、何か。日本語って曖昧な言語だとかよくいわれる。

斉藤：本当にそう思う？

江間：ものによっては、だから、分け方も違うわけじゃないですか。water がというふうなのと、お湯と水があったりだとか。そういう中で、内部処理とかいうのが言語、言語化っていう一つのモデルで考えるときに、何かそこの特性というか、何か一般モデルでいえるものなのか、もうちょっと文脈依存的なのか、どうなのかなとか思うんですけど、そういう話とかっていうのはあるんですか。言語……。

斉藤：え？　チョムスキー[47]がずっと言ってたゼネラルシンタックスみたいな。

大澤：汎用文法でしたっけ。

斉藤：あるかないかみたいな議論はもう決

[47] Avram Noam Chomsky. アメリカの言語学者・哲学者。生成文法の提唱者として知られる。

着がついたのかな。よく知らないけど。

大澤：負けた（　）（笑）。

斉藤：言語の差は、さっき言ったシリアライズする。要するにそうせざるを得ないよね。だから、語順があっちになったり、こっちになったり、いろいろするわけだけど、でも、それはその集団で、それがみんな、合意されてて、それこそコミュニケーションするために、それに沿わなくちゃいけないんだから、そっちにそろうんだと思うけど、もうちょっと前でシンボライザーとかっていうと単語だよね、ぱっと見て。「コップ」とか「パソコン」とか「マウス」とかって。そういうのはどの言語にもあるわけじゃない？　みんな、誰でも認識するよね、基本的なものとして。だから、それが、何ていうんだろう。

江間：基本的に、ここでいう言語って、単語レベルに落とし込むような感じなんですか。

斉藤：だって、それより下は何となく思い浮かばないよね。頭ん中でシンボルになってるっていったときのシンボルの第一候補は言語、単語……。

江間：何か関係性ではなくって？

斉藤：関係性じゃなくて、単語じゃないかな。だって、赤ん坊だって2語語。最初は1語で言いだして、「リンゴ」とか「まんま」とか「お母さん」とか言いだして、その次に2語語になるんだよね。だから。

大澤：何とか何とか、みたいな。

斉藤：そうそう。「パパ、行っちゃった」とか、なるんでしょ？　だから、それは関係とかってのは後から出てくるんじゃないかな。

大澤：シンボルは、そういう意味では、何か違うものがあっても同じ呼称でいえるかっていうのが一つの例えば定義（　）。

斉藤：ああ。個体のリンゴ、集合に名前が付いてるみたいな？

大澤：持ってきても。

斉藤：確かにそうだけど。集合とかいう概念は後付けだと思うな、私は。数学者が、数学ができた後で出てきたんじゃないかしら、みたいな。

駒谷：なので、今、シンボルの定義みたいな話になってると思うんですけど、どのシンボルに着地するかっていうのは、言語という意味ではかなり慣習があって、慣習に引きずられていて、何かディープニューラルネットワークとかで、何かシンボルが、最近の埋め込み表現とかそういうのがあるんですけど、そういうのって人間が理解できるシンボルとは違うシンボルが一番、説明可能性が高いみたいなものとして、今のニューラルネットとか持っちゃうんじゃないかなと思いまして。つまり、ニューラルネットが思ってるシンボルと、われわれが日本語というか、人間が使ってるシンボルっていうのがかみ合うかどうかっていうところが実は結構、難しいかもしれなくて、っていうか、かみ合えば、先生のおっしゃるとおり、昔のAIにきちんとつながって、そういうシンボル処理ができて、推論とかもできると思うんですけど、今、何かちょっとそこが、少なくともニューラルネットとかが統計

的につくり出しよるシンボルが、何かもしかしたら人間の思ってるシンボルとちゃうんかもしれへんなというのはちょっと思ってます。例として、さっき、江間さん、おっしゃったように、お湯と水が英語だと water だとか、日本語はやたらと何か……。ちゃう。英語だと例えば牛の体の部位とかがやたら細かいとか、それも、やっぱりそういうのを分ける必要があるから単語が違うわけで、それ、分ける必要があって、かつ、大澤さんが言ったように、インタラクションでお母さんが「いやいや、それ、同じリンゴやって」とか「これとこれ、同じ言葉、使たらあかんて」みたいなんを言って、やっぱり他人とインタラクションするのにシンボルが必要だっていうのはかなり本質的なとこかなと私も思ってまして。すいません。何が言いたい。何かあのモデル、僕、すごくいいと思うんですけど、ただ、その一番上のどこに着地するのかっていうところが、もしかしたら人間が思ってるのと、コンピューターがボトムアップに持ってくるもんで、もしかしたら違うかもしれへんなと、今、思っています。

江間：あと、すいません。ついでにっていうか、今、聞いてて思いだしたんですけど、単語で、物だったら分かるんですけど、概念とかはどうすんのかっていうと、すごく、それ、難しいと思ってて、今も AI とか、内閣府とかの会議とかでも、アカウンタビリティーというのに対応する日本語にはないみたいな話があって、アカウンタビリティー、今、説明責任って訳すけど、答責性と訳すのが正しいみたいなこと……。

斉藤：え、何性？

江間：答責、答える責任の性。

斉藤：答責性、初めて聞いた。

大澤：なるほど。そうですね。

江間：アカウンタビリティーっていうのは法哲学の先生とかに言わせると、何か問題が起きたら、もうそれは自分が腹を切るくらいの覚悟を持つっていう、そこくらいまでの意識というか、覚悟があるという。駄目だったら、もう金で払うと。そこまでの覚悟があるんだけど、日本で説明責任いうと、説明したらいいんでしょっていう、そういう話になってしまって。

大澤：確かに違う(笑)。

斉藤：重みが違う(笑)。

江間：そこの辺の概念が全くないのに、説明責任とか責任とかいわれて。だから、ロボットに説明責任があるのかとか、人工知能に説明責任あるのかとかいわれたときの何か求めるもの、機械に。その辺が全然違うって話があって、日本語と英語だけでもこんだけ難しくて、しかも、それはちゃんと、何か言葉が体に染み入って、その言葉を使えているわけではないのに、そこをどうやって……。別に批判したいわけじゃなくて、そこをどうやってすり合わせて、人同士でもすり合わせるの難しい中、人と機械とかのやりとりの中で「あんた、ちゃんと説明責任、それ、ちゃんと誓えますか」ぐらいな感じで意思疎通ができるようになるのって、

なるのかなってちょっと。

斉藤：いや、人間同士でも、誤解の山だよね、多分。

大澤：確かに。人間同士でできてないわけですから。

斉藤：だけど、だけど、だから、何、サイエンスっていうのは、一応、誤解の山みたいなのの中で実験をして、誰でも同じように試せて、再現できるとか、間違ってたら、またひっくり返せるとかっていう原理の下に事実を積み上げて、誰でも納得する。誰でももう、だから、物理とかそういうのの法則を否定する人はいないよね。否定したら、バカ（　）って感じになるわけじゃない？　そうやって積み上げてって、やっと対話が成り立つ。普通の日常生活で、常識的に身に付けた言葉でやりとりしてたら、もう誤解の山だよ、多分。結婚してる人は分かるけど、夫婦なんてもう（笑）。いや、本当にそう思わない？　何ていうのかな。でも、それは背景があるから、よくよく聞くと、その人はその言葉をここでこう覚えたとか、周りの人がこう言ってたとか、案外、理由はあったりして、それも聞き出すとだんだん分かってくるとかってなるよね。

江間：何か一方で、でも、何か誤解したままでも意思疎通ができたりするんだったら、それでもいいじゃないっていうような。ご結婚されてるような人もいる。何か、何とかなるんですか。

斉藤：いや、それはあり得ないんじゃないの？　だって、誤解してたら、どっかで破綻するもんね。

江間：同床異夢とか、呉越同舟とか、そういうような考え方ではなく？

斉藤：いやいや。だって、それはまずいじゃん、やっぱり。

江間：まずい？

斉藤：一生続けられると思うかって、多分、続かない。

江間：レベルにもよると思うんですけど、よく、よくじゃないですけど、海外の論文で「Butler's lie」とか、要は「執事のうそ」とか。「White Lie（白い嘘、罪のない嘘）」みたいなのの研究、一時期やってたんですけれども、大澤さんとこの人狼知能[48]とか、どうなのか分からないけど。

大澤：いや。すごくよく分かるのは、何かつまり、夫婦の場合は高度にあれですよね。意思疎通ができてないと、いろいろ不具合がありますけど。

一同：そうか？（笑）

大澤：疑問が出ている。かもしれないけど。

服部：単純にこっちが受け入れてるだけですよね。まあ、いいか。

江間：いや、でも、何か「Butler's lie」とかにやると、何かもめ事とか起きて、旦那さまが執事に「何か問題があったのか」って「いや、何もございません」みたいな。ホワイトライみたいな、うそも方便みたいなような考え方を機械がちゃんとインターフェース上に

---

[48] 人間と自然なコミュニケーションを取りながら、ゲーム「人狼」をプレイできるエージェントの構築を目的とした研究プロジェクト。

どう実装するかみたいなことを結構、2000 年代ぐらいで読んでたんですけど、私はその辺の論文を。

斉藤：へえ。

江間：結局、インタラクションの研究のほうになっちゃう。それは本当に先生がおっしゃるように、ちゃんと理解しているかっていうのとは別の問題だとは思うんですけど。

斉藤：でも、そのうそは、何かちゃんと理解してないとつけないんじゃないの？

江間：文脈を。

斉藤：相手、相手の意図してることとか。

大澤：そうですね。そこで意図のモデルが結構、大事になって。

江間：そこまで行くのか、インターフェース上で（　）。

斉藤：だって、表面的に調子合わせてたら、やっぱりどっかで破綻するような気がするけど。本当の意味で、相手のことを分かってないと。

大澤：だから、理解は必要なんですよね、多分、それをやる場合も。お互い違うモデルを持ったり、うそをついている場合でも。それは多分、ビジネスの場とかでは、誤解したなら、させたままでもいいかっていう市場はあり得るわけで（笑）、そこは興味がありますけど、そういう知能には興味がありますけど。そうですね。

駒谷：囲碁は常に盤面を共有してるから、やっぱり計算機と人間の間でかみ合ったというか、きちんと分かり合えた。分かり合えたっていうの、変ですね。きちんと話が、話が（　）。

斉藤：情報、完全情報ゲームですからね。

駒谷：そうです。その言語とかに、単語とかにグラウンディングしようと思うときに、だから、どれだけ、さっき言った計算機と人間の理解が違うかもしれないっていうのもあるし、人間同士も違うかもしれないっていうのがあって、何かちょっと囲碁の話と言語の話、人間の言語とか理解とかいう話を同じ土俵でできるのかなっていうのは、私、ちょっと違う（　）。

斉藤：それは違いますよ。だけど、囲碁はゲーム。ゲームはトイ、トイワールドなんだけど、だけど、言語屋さん、一生懸命、言語を研究してるけど、もうちょっと地に足が付いて、計算機で実装できるようなレベルまで理解したいって思ったら、スモールプロブレムから始めなくちゃいけないじゃん。そういう意味で。スモールだけど、案外、いろんな要素を含んでてっていうんで、私は囲碁は面白い対象だなと思ったんです。例えば囲碁用語っていっぱいあって。用語がいっぱいあるんですよ。分かってる人同士はその言葉で通じ合うんだけど、知らない人は分かんないんだけど、そういうレベルで話が通じるし、でも、さっき言った、そうでないレベルのパターン的な知識もいっぱいあるし、それが何か結び付いてるんだよね、頭の中でどうにかうまくね。だから、確かに本当の言語は難しいよ。今、AI スピーカーでしゃべれます、聞けます、何とかって言ってるけど、そんなレベルではないものがいっぱいあると思うけど、でも、その厚みがどれぐらい深いかは分かんな

いな。日常会話は案外、ええ加減にしゃべってる。

大澤：そうそう。それがつまり商用になるレベルじゃないと出てこないっていう難しさもありますよね、多分。その言語自体はかなり複雑な構造を持ってるけど、みんなが要するにちょっとスマートスピーカーに話すのは、結構、表層的なあれで済んじゃうかもしれないっていうような話は別にあるかもしれない（　）。

斉藤：日常的にはね。

大澤：どこが（　）分かんない。でも、そのシンボルの話でさっきあったのは、Google の AI がゴリラと黒人を見分けられないっていうんで、すごい話題になりましたけど、ああいう問題、確か、今でも解決してないので。

斉藤：解決してないか。

大澤：ええ。できないと思います、そんな簡単には。

江間：できなくて、結局。

大澤：つまり映像からは見分けが付かないので、それ以外の結局。

斉藤：だから、マルチモーダルにしないといけないってこと？

大澤：ただ、文脈のほうが……。

斉藤：実物だったら、実物で間違う人はいない。

大澤：いや、結構、難しい。

斉藤：実物で間違う人がいる？　人がっていうのは、人間ができてることかって聞いてるんだけど。

大澤：人間ができてない場合も。ぱっと見たら、分かんないかもしれないです。3 秒ぐらい。

斉藤：3 秒。

大澤：ぱっと見て、どっちですかって。

斉藤：だから、インタラクションを含めれば、できるようになるってことだよね。

大澤：可能性はあります。ただ……。

斉藤：だって、呼び掛けて、答えが。

大澤：そりゃそうですね（笑）。

斉藤：自然言語で返ってこなかったら、こいつは人間じゃねえだろうと。

大澤：Hello って言って。ちょっとそこも、ちょっと何で話し掛けるかとか。

江間：それが何か高度に行き過ぎて、プロファイリングは入ってくると、また別の問題が出てくるけどね。

大澤：うん。いや、つまり思ったよりもかなり文脈が、かなり依存するとこがあって、結局、Google は 1 回もミスれないって判断になって、結局、ゴリラってもの、そのものを認識しない形に、今、なってるみたいですけど。要は避けていると。だから、そういう問題も何かあるような気がしていてた。非常にこう、言語のシンボルつくるときに、かなり高度な推論が入ってるところはあるっていう。

斉藤：まあね。でも、人間も間違うと思うけどな。

大澤：間違う（　）。

斉藤：画像だけだったら、例えば。間違うよね。

大澤：そうだと思います。

斉藤：間違う確率が低いかもしれないけど。だから……。

大澤：人間は間違っても、その人、首にすればいい（笑）。それはあんまり、あれですね。まだ謝りゃいいってなるけど。

久木田：Ponanzaをつくった山本さん[49]っていう人の本に書いてあったんですけど、あれはやっぱりAIも間違うんだとか、そういうふうに解釈するのは全然間違いで、要するにあれは何かインターネット上のデータ、拾ってきて、黒人をゴリラって呼んでる人がいるから、そういうふうに覚えちゃったんだというふうに山本さんは書いてて、本当かどうか、僕は知らないんですけど。

大澤：なるほど。そうすると、さらにややこしいな。不適切な文脈で使ってるってことですよね。

久木田：そうです。

大澤：でも、不適切な文脈の中では正しい。正しいっていうのは、つまり倫理的に正しいんじゃなくて。

斉藤：意図的にそういう言葉遣いをする人がいる。

大澤：その中のコミュニケーションが成立してるっていう意味では理解はできるんですけど、それを他の文脈に当てはめてはいけないって問題ですよね。でも、その2つをただデータ取ってきて、区別できるかっていうとつらい。

服部：そうですね。

江間：そういった何かやっぱり人間のような知性を持っていったときに、それだって定義できないし固定できないし、時代によって変わってくものなんですかね。

斉藤：いや、あんまり進化のスピードと比べると、何か人間、そんなに進化、変化しないから、基本的な機能はある程度、定義できるんじゃないかと思うけど、普通に言葉を何万語ぐらいは知ってとか、普通に2カ国語ぐらいはしゃべれてとか。

大澤：Cycのプロジェクト[50]みたいな、つまり常識を入れるみたいなやつってのは、結局、どうなったのか（　）。

斉藤：Cyc、まだ続いてるみたいね、あれね。会社として何か、何をしてるか知らないけど。でも、最初は野心的だったけど、何にもなってないんじゃないかな。

大澤：あれも何か最初、白人男性を仮定してとかいう感じだったかなと思った（　）。

斉藤：いや、でも、そういうところの問題じゃないと思うよ、私。入れるべき。それは何か、そういうマイナーな問題もあるかもしれないけど、入れるべき知識の量がっていうか、言語化できてる知識だけを入れるわけじゃない？それこそ。だよね、あれね。

大澤：そうですね。

斉藤：だから、ここは何もないまま、それこそ判断をしたり何かする。推論する。シンボル、シンボルっていろいろな使われ方をするので、未定義語だけど、いや、計算機の中のシンボルっていうと、もうほとんど意味がくっついてないシンボルだよね。

大澤：そうですね。

斉藤：Identifierで、AとBは違うってい

---

[49] 山本一成。愛知学院大学特任准教授。将棋ソフトPonanzaの開発者。

[50] 一般常識をデータベース化し（知識ベース）、人間と同等の推論システムを構築することを目的とするプロジェクト

だけなので、こういうものは何もないから、計算機は。LISP で a って書いたら、それは単に a という変数なわけで、みたいな。だけど、人間がいろんな言葉を使うときには、もう過去の経験の全部の何か、いろんなものがその周辺にくっついてるわけだよね。っていうか、それをいつでも呼び出して使える、頭の中だと。ものを書くときでも、人の話を聞いたときでも、ぱっと何か見たときでも、過去の経験を全部、呼び出して、それを使って、何か判断できるよね。そういうとこが随分違うと思うんだよね。それを、だけど、つくるのは大変難しくて。

大澤：(　) そこがあって。何かちょっと前に……。すいません。ルーク・スティールズ[51]、人工生命の言語をやってる方が、何かシンボルグランディング問題、解決したっていうのをもう書いてて[52]、読んで、松原先生とかといろいろ議論をして、「私はこれを認めない」みたいなのが面白かったんですが、ロボット同士でそういう言語生成をやって、あるシンボルを持たせることはできるんですよね。確かに、そういう意味でルーク・スティールズの言ってること、間違ってはないんですけど、ただ、それは人間が思っているシンボルとはかなり差があると。だから、環境とか何かのタスクっていうのは、実は結構、重要なのかなと個人的には思ってるんですが。置かれた環境が非常にシンプルな課題しかやらない、シンプルなロボットはセンシングしたものに対して共通の情報を理解して、あるシグナルをつくって、それを他者と共有するっていうのは十分できるわけですけど、それが人間にとって意味のあるシグナルであるかはちょっと分からないっていう。

斉藤：まずさっき、駒谷さんも言ってたように、何か人によって、みんな違う。

大澤：定義が、だから、違うので、これは、だから、人工生命屋さんらしい定義なのかなと。

斉藤：でも、お互いに話が通じるときには、大体、大体合ってんだよな。

大澤：大体合ってます（笑）。

斉藤：いや、それは色に名前がいろいろありますとかって言うけど、だけど、大体合ってるわけじゃない？　赤と緑が同じ色に見える人はいるのか。分かんない（笑）。いるけど、そういう例外だからさ。大体、スペクトラム上でここら辺が赤、ここら辺が緑みたいなのは、大体、合ってるわけじゃん。それ、人間のメカニズムが、目とか、視神経とか、何かそういうところが共通だからだよね。

大澤：そうですね。あと、ただ、文脈の影響もあると。

斉藤：文脈の影響もある。

大澤：最近、自動運転の話で、ブルックス[53]が言っていて、結局、赤い停止のあれを見ても、実際の画素は赤くはないと

---

[51] Luc Steels. 人工生命研究者。

[52] The symbol grounding problem has been solved. so what's next, Luc Steels, https://philpapers.org/rec/STETSG

[53] Rodney Allen Brooks. アメリカのロボット工学者。

いう。当たり前ですけれども。だから、周囲の環境光とかと比較して見るわけですよね、人間。だから、要するに本当に赤を定義するのは、実はまだできてないというような話で、ちょうどそういう話をして。だから、思ったよりも結構……。つまり、そのレベルのAI もまだできてないっていう意味では、そうかもしれないです。

久木田：何年か前に、確か、阪大にそのスティールズさんが来て、講演してった。私、聞きに行ったことがあるんですけど、面白かったんですけど、やっぱり今、言ったようなアプローチで、こうロボット同士がコミュニケーション取って、例えば「あれを持ってこい」みたいなことを言って、それをちゃんと持ってきたら、ちゃんとグラウンディングできてるっていうような話だったんですけど、そのときに質問した人が、例えば「ボトルを持ってこい」って言って、持ってきたとしても、例えば僕らみたいに、ボトルってのが水を入れるもんで、水っていうのは飲むもんでみたいな、そういう経験を欠いていたら、それはちゃんとグラウンディングできてないんじゃないかみたいなことを言うと、その人は、いや、そんなことはない。シンボルのグラウンディングっていうのは、あくまで一つの言語ゲーム。彼、だから、ランゲージゲームっていうふうに呼んでますよね。言語ゲームの中で何かが果たせると。役割を果たせるっていうことであって、それは何かいろんな意味付けをやってたら、無限にたくさんの情報がそこに込められてなきゃいけなくなるけど、そういうもんじゃなくて、限定された状況であることができたら、それで。

大澤：だって、それはそうです。だって、そのロボット、水飲むわけじゃない(笑)。グラウンディング作用はどうでもいいじゃないですか。それは全て人間中心的なあれですよ。

久木田：そうですね。人間つくらなきゃ。

大澤：そう。ボトルじゃないと認めないってのは。

久木田：教わらないってことになりますかね。

大澤：そうそう。分かりません、ロボットの気持ちが分かりません（笑)。

江間：そこまで期待されてもって。

大澤：そう。すいません。

江間：何か、話がちょっと戻ってしまうか、先、進むのか、分からないんですけど、要は、先生はそういうものをつくりたいって思うことを言語化していただくことっていうのは可能なんですか。何で、何でそういう強い人工知能みたいなのとかにご関心が？

斉藤：だって、単純に不思議だよね、人間のやってることって。子どもを見てるときに特に強く感じるけど、例えば何だろう。言語爆発みたいな現象とかさ。ほとんど 1 回言っただけで覚えちゃうよね。計算機は今のディープラーニングでも何でも、何万、何千万回とかって繰り返して、やっとパラメータがチューニングできる。でも、人間の場合は何かあっという間に覚えちゃうんで、私はそれは言語をうまく使って

るからだと思うんだけど、お母さんが子どもに言う言葉とか何だとか。だけど、不思議過ぎるよね、人間が脳でやれてることって。ものを理解するんでもそうだし、問題を解くんでもそうだけど、普通にしゃべるんでもそうだし、言語を理解するのもそうだけど、あらゆることが何か不思議過ぎるから。で、どういう仕組みなんだろうって思うよね。私は仕組み、仕組みにとっても興味があるので、子どものときから時計壊したりなんか、ずっとしてたので、メカニズムっていうのにすごく興味があって、それを知りたいって思うから、この脳の仕組みはどうなってんだろうって純粋に思うよね。それのいい手段だと思って、AI に飛びついたけど、なかなか何にもできないな、みたい（笑）。認知科学にもちょっと入ってみたけど、認知科学も何か、ええ？最近、認知科学は久しぶりに行ったけど、ほとんど何か統計処理だけやって、「あ、有意な差がありました」って。

大澤：だいぶ心理学に影響されてるような。

斉藤：あれは心理学が悪いのかもしれないけど、でも、じゃあ、それで、「それで何なの？」って聞くと、何か黙っちゃうみたいなさ。何が知りたかったの、最初に、みたいな、そういう研究が多かったりして。っていうか、認知科学、頑張ってると思うけど、だから、問題が大き過ぎるんだよ、多分。脳の仕組みを知ろうという問題が大き過ぎて、私は若気の至りでそれを AI でできるかもしれないとかって思って、やりだしたけど、全然できなかったし、認知科学も少しは勉強したけど、でも、まだ全然分かってないな、みたいな感じだし。でも、脳科学なんか頑張ってるかもしれないけど、でも、それも随分遠いよね。

大澤：そうですね。

斉藤：と思う。だから、でも、いろんな学問分野の人がいろいろやってくれて、いろんなことが分かってくると、だんだんにその全体像みたいなのが分かってくるかもしれない、知能に関して。単純にもう不思議過ぎて、知りたいよねっていう純粋な疑問なんだけど。

服部：かなりあれですよね。学問的なところですね。それが、そういう興味がドライブしてるような、それが産業になりますかとか、そういうこととかでは全然なくて。

江間：サイエンスとしてって、人間理解のためですよね。

斉藤：そうだね。だって、いや、もう何か、最近、ノーベル賞取った人が、みんな、話するから、古びた話かもしれないけど、結局、世の中に役立つかどうかはその後の話で、だって、本当に分かったら、随分、いろんなことで役立つよね。教育なんて、教育の分野って、みんな、一生懸命やってるけど、何とかラーニング、何とかラーニングとかってさ。でも、結局、その学生 1 人が頭ん中で何が起こってるかが何も知らずにやってるわけじゃない？そんなの、何かあれだよね。うまくいくはずないじゃん？　だから、本当は何か学習してるときって、何が起こってててとかっていうのをちゃんと分かった上

で、そこにぴったりな例えばアドバイスを与えるとか、何か教材を与えるとかってうまくやってけば、だから、実験的なそういう研究はいっぱいあるから、それでもいいのかもしれないけど、でも、最終的な応用は、もし本当に分かったら、もう何か……。分かったら、つまんないってなっちゃうかもしれない（笑）。あ、何だ。人間ってこんな、こんなばかな仕組みだったのか。みんな、すごい、すごいって言ってたけど、脳って何か、こんなばかなことやってたのかってなるかもしれないけど。

服部：何か今の教育の話だけじゃ、多分、ないんだと思うんですけど、ああいう、言ってる、世間でやってるのは、大体、あそこの図で言うとこの言語観。上のちっちゃい三角形のところ。だから、言い方ですよねとか説明の仕方、言葉のつむぎ方ですよねっていうところで、みんな、今、一生懸命やってるんだけど、本当はその人の何かよく分かんない、何かやってると。何か「分かった」っていうふうに思うものって、本当は何なのっていうのが分からなければいけないんだけれども、そういうとこに全然、こう、アタックしていくような研究って、まだまだ全然、各分野、どの分野を見ても AI は、そういうところが何かこう、分かるようにするのかなと思ったけども、そういうところにやっぱり行かないし。

斉藤：すぐには行かないってことだよね、多分ね。でも、またブームが去り、次のブームが来て、こうやって何度もやってると、だんだんにいろんなものが分かってきて、いくかもしれないとは思いますけど。

神崎：いや、一応、英語教育とか、割と最近、ニューロの人が増えているというか、最近、雇われてる人は大体、そういう人じゃないかなという。

斉藤：英語教育。

神崎：要するに英語学習時に脳内で何が起こってるかっていうのは。

斉藤：分かってんの？

神崎：いや、測っている。そういう系統の仕事が増えてはいる。

斉藤：増えている。

駒谷：それはこの辺から脳波が出るとか、そういう話ですか。

神崎：例えばどういうときに、ただ、どういう形の設問とかで、ある部分が活発に動いてるかとかを見ることによって、前より効果的なほうを見るとかっていうのは教育の分野ではあります。

大澤：頑張って理論化しようとしているっていう。

神崎：ただ、それも仕組みが分かってるわけじゃなく（　）ボタンが分かるけど、中身が分かんないみたいな、そういう感じのもんだとは思う。

江間：分かっても、別にニューロフィードバックしてもいいかはまた別の話だから。

神崎：そこまでではない。

大澤：問題変えたりとか。

駒谷：結局、それもこういう入力データはこの辺がピッて光るみたいな話で、何か入出力の関係を取ってるだけなんだったら、ディープニューラルネットワ

ークと一緒じゃない？（笑）

大澤：中で何やってるか分からない。

服部：そういうのが分か……。そういうのは何か計測して分かるよねっていうところと、何か言葉で、こういうふうだよねって説明できるというところの間の、ここの隔絶したところがあるっていうことなんですよね。

斉藤：うん、分かってないところ、分かってないところがある。それで、下から一生懸命やってる。だから、ビジョンの研究とか、第 1 次視覚野第 2 次視覚野とかっていって、少し分かってきて、それこそ Marr とかが始めた研究の方向で、ボトムアップに、最初はこういう情報を抽出してるんだとか、それを組み合わせて、例えば形、3D の形が認識できるようになるんだ、みたいな、下からこう攻めてる人たちもいるし、でも、それって最後には、何かもっと複雑なものを見て、さっとそれの意味を理解したり、また文字なんかそうだけど、文字を読んで、中身を理解してるよね。

江間：本当にしてるかどうか（笑）。

斉藤：ほとんど日常的に文字を読んで、書いてあることを理解してるよね。それも何かすごいことだよねっていうか、単にドットの（笑）。

大澤：そうですね。

斉藤：音声もそうだし。音声なんかも。

大澤：確かにそう。何か、だから、言語の線状性なんか、この前、手話やってる人とその話をちょっと聞いたんですが、要するに手話だと、シンボルって別々に表すことができるので、例えば。

斉藤：2 つ 3 つってこと？

大澤：そうですね。顔の動きと、左手、右。そうすると、同時にぱっと出るわけですけど、そういうことって、実は言語ではできない。

斉藤：できない。

大澤：どうしても線状に並べないといけない。そこで文法の違いとかが出てくるのと、割とやっぱり思ったよりもかなり高度なことをやっているというのは間違いないのかなっていう気はしますね。

斉藤：しかも、ほとんど努力なくっていうか、自然にできちゃう。

大澤：努力はしたような気が（笑）。

斉藤：それでも、子どものときに努力した記憶ないよね。

大澤：ああ、なるほど。母語はそうですね。

斉藤：言葉を覚える最中に努力した記憶は全然なく、知らない間にしゃべれちゃったよね、みたいな。

大澤：何か注意されたりっていうのは。

斉藤：自転車に乗るときは随分、努力したけど、自転車に乗れちゃったら、もう全然、平気だけど、そういうのとは違って、言葉は何か、あっという間に、みんな、知らないうちに。

大澤：何でしょうね。自転車はそう。失敗したときのダメージも大きいから、結構覚えてるって。

服部：何か子どもを見てて、これは多分、周りが言うからだと思ってんですけど、何か「僕、頑張った」って言ってきたときに、頑張ったって、何をもって頑張ったって、こいつは言うのかっていうの、あんまりよく分かんない。多分、

多分、でも、きっと幼稚園で何か「頑張ったね」って言われたから「僕は頑張った」っていうふうに言ってるんだなと思うんだけれども、ひょっとして、何か自分がああやって、こうやって、こうやったって何か、ごにゃごにゃってあって、それを何かのところで「頑張った」っていうところにぽこって、そういう言葉で表すっていうのを、何か知らないけど、するようになってるんだとすると、それは一体、何なんやろうねとかっていうのは。

斉藤：それはちゃんとつながって。生理的なものとつながってると思うけど。汗かいた、みたいな。走った後は「頑張った」って言うよね。徒競走、運動会でぐるっと周って走ってきたら、もうはあはあ言ってるわけだから、物理的にやっぱり体がダメージっていうか、ストレスを受けてるじゃん。その情報とかは当然、どっかでつながってて、「頑張った」っていう言葉に。

服部：ストレスか。何かすごい複雑な絵を描いてくるんですね。何かこうあって、すごく細かく色を分けてるわけですよ。「ああ、よく描いたね」とか言うと、「僕、頑張った」って。

斉藤：頑張った。時間とか。

大澤：いろんな違う現象に「頑張った」ってやっぱり言うわけですよね、それは。

斉藤：言葉はすごいよね。

大澤：そこは何か共通の概念をやっぱり。

服部：それ、聞いてると、だんだん、そのうち「うーん、分かんない」とかって、大体、逃げるわけですよね。確かに分かってないと。そうかって。

大澤：説明は難しいんでしょうね。人間だって難しい。

久木田：「生意気」とかいう言葉について、「何とかは生意気だな」みたいなことをいうときに、「生意気ってどういう意味？」って。

大澤：哲学者だあ

久木田：「それは分からない」（笑）。それは分からない。

服部：そういうの、ちょいちょいありますよね。「それって、何でそう思ったの？」とか「それってなあに？」って言うと、「うーん」ってなるんだけど、だけど、勝手に言ってるわけですよね。そういう言葉を使って、何かそういうとき、まあ、悪くはないなっていうタイミングでそういうのをぽんぽんと出てくると。それが何かものの名前とかじゃなくて、そういう何か概念みたいなのを出してきたときに、うーんってやっぱり思うわけですよね。それってっていうのはやっぱり、ちょっとうずくものはありますよね。それってな、みたいな。どういう情報処理かなとかっていうのを。

久木田：不思議ですよね。何かそういう物理的な何か感覚できないものにちゃんと名前を。僕、一番不思議に思ったのは、結構、2歳ぐらいのときに、もう既に「全部」とか、「ない」とか、「かもしれない」とか、そういう何かロジカルなワードがでてくる。

服部：それ、何か不思議な、不思議に使うときってありますよね。そういうの出してくる？みたいな。それがどういうメカニズムかっていうのはよく分かん

ない。でも、出してくるねとか。

久木田：それこそ、チョムスキーを信じたくなりますね（笑）。

大澤：だから、速いなって、学習が（笑）。生成文法。

服部：それが、僕ら、それ、さっきの何か難しいとかっていう言葉になってくると、はあってなるんだけど、例えばコンピューターに何かやってるうちに、そういう言葉、あのちっちゃい三角形の線の上ですよね。その下のほうから上に上がってこなくて、コンピューターは何か、僕らが分かる言葉じゃないけれども、それって何だっていうのは何か分かってると。何かの計算で何か分かってるっていうのは、例えば分かったとして、でも、それだとわれわれとしては十分じゃないわけですね。何か何が、われわれの知性っていうのはこういうメカニズムなんだっていうのは、僕らには分からないわけですよね。そうすると、やらなきゃいけないことは「あそこの中、分かんないね」っていう、下の台形のところと、それをいかに言語に結び付けるかっていう、そういうところもやっていかないと、最終的には知能っていうものがこれだって分かったっていうところにつながらない。だから、非常に距離が遠いねっていうことですよね。下の台形のところだけだと、あれですよね。何か医者とかだと、何かこう、普通、診察とかして、いろいろセンシングすると。普通の人間の医者だと「これこれこういう症状。だから、あなたはこういう病気ですね」って言うんだけれども、コンピューターがもしも何……。コンピューターはそういうことしないかもしれないですね。何か「こうこうこうだから、あなた、こういうふうに人生を送ると、一番、長生きできます」とかって、もはや病名とかの言語にはしない（笑）。何かそういう、何かよく分かんないけど、こういうもの、こういうふうにコンピューターつくったら、何か人間を長生きさしてくれるっていうことは分かるんだけれども、それが何かどういうふうな症状だというふうにコンピューターは判断して従って、こういうふうに推論した結果、こういうふうに生きて、生活しなさい」っていうふうに言ってるか、僕ら、分からない。分からないけれども、長生きできるのね。まあ、信じましょう、みたいな。

江間：うん。お告げですよね（笑）。

斉藤：え？

江間：お告げ。

斉藤：お告げ。でも、それをつくるのには、何か医学でもやってる、生活習慣みたいなデータ、山のように取って、それ、ぶち込んで、計算機に。学習させて、こういう生活パターンを取ると一番長生きできますっていう結論が得られてるんなら、名前が付いてようと、付いてまいと「あなた、ここではこうしなさい」っていうのはできるかもしれないよね。

服部：そのときに、人間の医者が「何でこいつはこういうことやったんだ。説明せよ」って言っても、そもそも病名にするっていうのはもともと機能として

持ち合わせていない。こうすればいいってことは分かってる。そうして、この人間っていう生命体っていうのは、より長く生きられるけれども、僕らが言葉として、取りあえず医学っていうところで、言葉、つくった言葉っていうのは全然、理解してなくて、そういう意味では、説明はコンピューターはしてくれない。だから、僕ら、分からないと思うんだけれども、だけど、とっても有益だっていうのになるわけ。

斉藤：まあ、なるかもしれない。

服部：ええ。でも、それだと、われわれはその知能というものをいかに今の医学、医療の分野で、コンピューターっていうのはどういうふうに機能してるか分かってっていうことにはならないですね。

斉藤：うん。まあ、そうね。今の例は、だけど、何か説明のしようはあるような気もしなくもないけどな。いや、くべた、くべたデータがどういうものかによるけど。何か囲碁の例は、本当に盤面のパターンだけなので、入力が。入力なり、中でリプリゼント（　）してるもの。本当にそれだけなので、余計な概念をつくらないと人間同士は対話できなかったし、計算機も何か、結論出してくれても、そこに何か人間と同じものかどうか分かんないけど、そういうのを見つけ出さないと。そういう研究もあるんでね、何か。ニューラルネット、つくった後で、ニューラルネット自体を分析して、何か概念みたいなもの、あるかないかとかって、抽出しようとかって研究もあって、そこで違うものが抽出されるかもしれないけど、でも、それを使って説明しだせば、説明はできなくないかもしれない。だから、それって全然、私が言ってるAI、元のAIとは違うものだよっていうか、ニューラルネットの説明責任を果たさせるために別の工夫をして、ニューラルネットの仕組みを何か人間が理解できる言葉に翻訳する翻訳機をつくりましょうっていう話だよね。それはできるかもしれないけど、やることの意義があるかもしれないけど、元の言ってるAIとは違う。

江間：翻訳機が後付けでできたとして、今の点も解釈して、するわけじゃないですか。それを埋め込んで、何か別のものがつくれたら、じゃあ、それはできたことになるんですかね。それは別？

斉藤：だって、人間がやってることと同じかどうかっていう、そこ。つまり人間の脳の中では、何をどうやってるんだっていうのを知りたいわけで、じゃあ、つまりそうやって埋め込んだのは同じかどうか分かんないよね。

江間：人間が分からないからっていうことですよね、結局のところ。

斉藤：そうそう。だから、人間、分かるほうが先だろうって思う、順序としては。人工知能って、最初、本当にナイーブに、面白いから、いろんなプログラムつくって、こうやってんだろ、ああやってんだろって言ってつくってみたけど、で、ある程度はできたりしたけど、でも、「それって本当に人間がやってること？」って問われたら「そうです」って言えないよね。何かそうい

う手法がない、手段がない。人間の脳の中でやってることと、つくったプログラムとの equivalence を証明する手段がないよね。だから、人間のほう、もっと調べて、こうやってるんだってのが分かるか、そういう何、対応をつける手段みたいなのが見つかるか、何かブレークスルーがないと。単にプログラムつくって、うまくいきました、は工学的にはいいんだよ。だって、今だって、もうニューラルネットでいろんなことができるようになって、それ、全然、いいことなんだけど、でも、中身を知りたいっていう目的に対しては何か進まないんだよね。難しい。その難しさにずっと、私はもう十何年前にその研究も離れて、事業部とか行っちゃったんだけど、でも、その当時からまだ何も。分かんないな。ちゃんと勉強してないから、少しずつ分かってきてるかもしれないけど、何も分かってないと思う、やっぱり。

大澤：認知科学はそういうポリシーですよね。つまりその中で何をやってるか。取りあえずモデルを立てて。

斉藤：最初の頃、そういう議論してたけど、もう今の認知科学は何か現象、「けんちゃんの絵日記」になっちゃってる。「けんちゃんの絵日記」って、研究、ある研究を非難、批判するのに使う言い方なんだけど、何か「今日は朝、起きました。ご飯食べました。外、遊びに行って、夜、寝ました」みたいなのが延々続いてるみたいな（笑）。

大澤：なるほど（笑）。

斉藤：何ていうの。現象を見て、その現象を記述して「こんな面白い現象があります。おしまい」みたいな。だから、本当に中でやってることを解明して、しようとか、したぞとか、ここはこうなってるんだ。もっと前だと、すごくラフだけど、short term memory と何とかがありますとかさ。そのモデルは、実際にも何か脳の調べたのって、ちゃんと証明されてるかどうか知らないけど。されてないのかな。でも、short term みたいなものがありとか、そういうのはある程度、ね。だから、そういう成果って、何か前に進んでるような気がするんだけど、一つ一つの何か面白い現象捉えて「こういうことで、こうやるとこうなりました。おしまい」っていうのだと何かつまんないよ。っていうか、学問、科学としては先、進んでないような気がするので。いいけど、いいけどっていうか（笑）。それで論文書けて、はい。就職できれば、それでいいんでしょうけど。

大澤：斉藤先生はかなり早い段階で、安西先生[54]の認知の learning by doing の話があったりしたので、認知科学の興味は結構早かったんですか、他の方と（　）。

斉藤：野田セミナーとか行ったし。

大澤：じゃあ、もう AI から？

斉藤：でも、認知科学って、いつできたのかな。

江間：87 年。土屋先生がそれこそあれですよね。創業の署名か何かに書いたとか

[54] 安西祐一郎。慶應義塾大学理工学部名誉教授。情報工学、認知科学。

いう話を。

大澤：ああ。日本認知科学会ですよね。

斉藤：日本認知科学会はもっと後だから。

江間：認知科学っていうジャンル、分野が？

大澤：認知科学自体は。

斉藤：分野ができた。

大澤：でも、結構、人工知能と同じぐらいというふうに、あれしてるんですけど。

江間：ほぼ似たような。

斉藤：人工知能と同じぐらい。いや、その本の頃には認知科学ってのはないと。

大澤：ないですか。サイモンとかあの辺の方なんですかね、言い始めたの。

斉藤：サイモン？　ああ、サイモン、ニューウェル。

大澤：うん。

斉藤：そうね。サイモン、ニューウェルと、そこにいた連中と少し動きが変わって。

江間：「ダートマス会議を認知科学の発祥と見る向きも多い」とウィキペディアは（笑）。

斉藤：も、多い。

大澤：ウィキペディアだからなって（笑）。

斉藤：まあ、そうかもしれない。だから、その頃、やっぱり似たような興味で。サイモン、ニューウェルは実験したりしたからね。クリプトアリスメティック（虫食い算）でも何でも、たくさんの被験者にやらせて、チェスもそうなんだけど。被験者のプロトコル取ったりしたから、どちらかっていうと認知科学的な方向だよね。あの頃、分かれて出てきたのかもしれない。

大澤：なるほど。

斉藤：その頃から、両方、興味あったけど。そう。どっちも。どっちもどっちだな。何ていうか、そんなこと言ってるけど。

大澤：そうね。読まれてる論文とか見ると、ちょっとずつカラーがやっぱ違うというか。中島先生はやっぱプログラミングの話が。

斉藤：中島秀〜、あいつ、アーキテクトなんだよな、何となく、私の……。何だっけ。彼、書いたじゃん。『知能の物語』[55]とかっていう。

大澤：はい、あります。

斉藤：あの本を読んだときにすごく感じたのは、彼はアーキテクトなんだよ、やっぱり。アーキテクトっていうのは、アーキテクチャーがつくれると万歳でおしまい。

大澤：なるほど。

斉藤：アーキテクト、アーキテクチャーと私が知りたいと思ってる細かいメカニズムとはちょっと階層が違うんだよね。アーキテクトっていうと、私、プログラミング、ソフトウェア開発の会社にもいたことあるんで、ソフトウェアのアナロジーが一番分かりやすいんだけど、アーキテクトって一番トップレベルで、お客さんと話をして「あなた、やりたいこと、何ですか」と。「それのためにはこういうデザインがいいですね」って言って、ソフトウエアの全体のデザインを決める人がアーキテクト。

大澤：なるほど、なるほど。

---

[55] 中島秀之, 2015,『知能の物語』, 公立はこだて未来大学出版会

斉藤：それは細かなコーディングレベルとか。
江間：現場レベルの話では。
斉藤：現場の細部はどうでもいいんだよ。
大澤：なるほど。
斉藤：この問題はこういうストラクチャーで解けば解けるぞとかっていうことが分かると、で、もう上がりになっちゃうような感覚があって、そういうことを一番目標にしてる感じ。中島秀之は。プロローグとかやりだしたら、もうそうだよね。あんなの、使い物にならない言語なんだけど、どう考えても。
大澤：なるほど（笑）。
斉藤：いや、だって、実際、誰もあんまり使ってないよね。
大澤：まあ、そうですね。
斉藤：だけど、それに飛びついたり、いいなと思ったりっていうのは、アーキテクトだからなんだよね。
大澤：原理的にはできる。
斉藤：いや、原理的にはできるけど、現実的にはできないよ（笑）。
大澤：現実にはあんまり、そこまでは業務。
斉藤：だから、知能に関しても、知能のアーキテクチャーみたいなのが分かればいいやと。それってかなりトップレベルの構造みたいなところ。でも、私も何か神経レベルにまで下がってどうしようもないんだけど、でも、もっと下のレベルで、頭ん中で、何と何がどうなって、つながってとか、そういうとこが知りたい。手段がないんだよね、結局ね。でも、最近、脳の研究も随分、進んできたから。
大澤：そうですね。部位レベルではかなり測定できるようになってきたんで、課題与えたときに。
斉藤：でも、でも、やっぱり何か、群盲象を撫でる以外の何も……。
大澤：部位レベルのあれじゃないですかね。ロジカルな限界はあるというか、同じとこで違う働きしてても分からないこと。
斉藤：でも、最近、そこに、そこに、そこに。そこのジュリストの上に置いてあるさ、そのペーパー。
江間：なぜジュリストが（笑）。
斉藤：それそれ。今度、読もうと思うんだけど。
大澤：ああ、なるほど。面白いですね。
斉藤：こういうの、何かどんどん出てきたんだよね。
大澤：じゃあ、かなり細かい精度でやってるとこ。
斉藤：いやいや、まだちゃんと読んでないから分かんないけど、そういうの、いっぱい出てきて、今日、読んだのは何か、人間は顔、人の顔を 5,000 人分、覚えてます、みたいな研究結果が出てきたり（笑）。本当かい。だから、シンボライザーって意味では、一人一人の人のシンボライザーがあるんだよね。そうだよね、多分。こういうような人。
江間：おばあちゃん細胞みたいな。
斉藤：おばあちゃん細胞もそうだけど。おばあちゃんってのはちょっと抽象的な概念だけど、でも、最初にできるおばあちゃん細胞っていうか、おばあちゃんシンボライザーは自分のおばあちゃんだよね、多分ね。でも、それがある程度、一般化されて、どんなおばあ

ちゃんにでも適応できるし、でも、おばあちゃん、いっぱいいるし、友達もいっぱいいるし。人間の顔のシンボライザーはそれこそ 5,000 ぐらいはある。この数、私、すごく興味があって、数のほうから押さえると、人間の脳ではそれぐらいできないやとか、単語が、人間の使える単語は 3 万語だとか、そういうのとか、そういうの、すごく興味があって、それが人間の限界なんだよね、多分ね。でも、囲碁のプレーヤーなんかはいろんなパターンのすごい数を覚えてるからね。それは記憶力、すごいんだよね、やっぱりね。人によって違うけど。

大澤：最終的には、それはつまりニューロンの、1 個のニューロンレベルは落ちる。

斉藤：1 個まで行かないと思うけど、最近の研究はみんな、コラムみたいな、何か複数のニューロンで 1 つのことをコーディングしてるってのが大体、通説だよね。

大澤：なるほど。

斉藤：だから、1 個だから 100、何、1,000 億個、ニューロンがあっても、1,000 億個の違うものを覚えてるわけでは全然ないはずで、そうじゃないけど、100 個ぐらい 1 束になって、1 つのことを覚えてるとか、1 つのシンボルに対応してるとか、本当にそれが正しければ、数……。

大澤：なるほど。

斉藤：数が一番、単純だと私は思って、数でまずは研究するのがいいかなと思ったんだけど、数自体がもう複雑だよね。

大澤：それでも、あれですね。やっぱ点が幾つあるかと。

斉藤：それ、subitizing と同じ実験をやってるんだよね。subitizing って知ってるよね。ぱっと見ると、さっと。subitizing ってすごく有名な話で、画像を見せて、その中に数が少ない間は……。っていうか、横軸にその数を取って、ドットの数ね。で、縦軸に反応時間を取ると、こうなるんだよ。ここが 3 とか 5 とかまでは、もう。要するに一瞬にして、パターンで認識して、数が分かる。でも、ここからは数えなくちゃいけないんで、カウンティングになって、もっとたくさんになったりして、estimation っていって。何ていうの。

江間：大ざっぱに。

斉藤：大ざっぱに 50 個とかって言いだすんだけど。人間のものを数えるのの subitizing っていうんだけど、それの実験と同じ実験をやる。やりながら、脳のあれを、反応を取って見てて、面白そうなんだけど。

大澤：個数とナンバーで比較して、やってんのかな。ぱっと見て。

斉藤：分かんない。でも、そういうの、どんどん出てきたから、脳のほうの研究から何か分かってくることもあるかもしれない。そういうの、うまく組み合わせれば。

大澤：それは設計的なとこ、追い詰めてくやり方ですね、ある種。

斉藤：そうだね。

大澤：この中で動いてることは間違いないだろうからっていう話ですね。

斉藤：そうだよね。あれ、だから、そう、そ

う。外のインタラクション、大事とか、ソーシャル、社会性とかっていうけど、まあまあ、元はこん中だから（笑）。ここ、何かブラックボックスにしちゃって、やったら、何かラフなセオリーはできるかもしれないけど、何かちょっと地に足が付かないよなって気がしなくもない。

大澤：どこを目指すかなと思うんですが。まあ、そうですね。なるほどね。でも、結構、何かしっくり来たというか、中島先生とかも認知科学会来て、結局、認知科学と人工知能学会、一緒のようなもんだから、みたいな話を、雑とか思ったり（笑）。でも、まあ、徐々にそういう感じ。多分、そこまで分化していないというか、多分、皆さん、興味というか、結構。

斉藤：違う人が入ってるし、違う学会だと思うけど。

大澤：今はそうだと思いますよ（笑）。思いますけど。

斉藤：うん。だから、知りたいことの大本は同じだってことだよね。でも、そこ、アプローチはいっぱいあって、心理学だってそうだろうし、もう個々のビジョン研究してる人もそうだろうし、音声研究してる人もそうだろうけど、みんな、そうだけど、何か。そうね。人工知能学会、だけど、できたとき、私はやっぱり何か違うなと思って、入らなかったんだよね。

大澤：そうですか。

斉藤：入らなかった。私、私は情報処理学会と、認知科学会しか入ってなくて。あ、ソフトウエア科学会に入ってたけど、もうやめちゃったけど。

江間：じゃ、結局、人工知能学会、1回も入られて？

斉藤：1回も入ってない。

江間：行ったことは？

斉藤：行ったことは……。もないかもしれない。

江間：そうなんですか。

斉藤：うん。すいません。ごめんなさい。

大澤：いいですけど（笑）。

江間：でも、中島先生とかとは、AIUEOのメンバーとは、今でも結構、会ったりとか話したりする場はあるんですか。

斉藤：あのときのメンバー、別のグループもあったりして、時々会いますけど。

江間：別のグループっていうのは？　。

斉藤：別のグループっていうのは、同期、同期のグループだね。

江間：同期？

斉藤：研究会じゃない、単に。

江間：コミュニティーとしての？

斉藤：コミュニティー？　何か分かんない。飲み仲間みたいな（笑）。

大澤：それはありますね。なるほど。

江間：先生も何か、SNSとかされてたりするんですか。

斉藤：SNS、してない。私はウェブページもないし、SNSもしてないし。何か忙しくて、そういうの。ウェブは自分でつくろうと、あるとき、思ったけど、時間がなくて、つくれてないまま、何か十数年、過ぎ。

大澤：結構大変ですからね、あれ。メンテナンスし続けるの。

斉藤：いや、でも、最近は何か簡単にできそうじゃん。ワードプレスとか、何ちゃ

らかんちゃらとか。

大澤：最近はあれですよ。科研に申し込むときはリサーチマップ入れなきゃない。

斉藤：リサーチマップ。

大澤：そう。つくらされるんですけど、何か自己紹介ページを。

江間：それは何か、そういう当時の仲間とかといろいろやりとりするのは、もっぱらメールみたいな、そういう感じですか。

斉藤：メール。そうね。メール。

江間：結構、土屋先生とかは、もう Facebook 上でいろいろ、堀先生と土屋先生がディスカッションを。

服部：土屋、堀、中川。

江間：あたりが。しかも、他の人の Facebook のページで議論を始める、みたいな。

斉藤：ああ、中川、中川、中川。中川氏も、そうね。よくやってるね(笑)。いや、私は何か、このちいちゃい画面にいっぱい書く気はしなく。まだメールのほうが、メール人間だから、メールのほうが多いけど、メールでっても、そんなにディスカッションにはならないかもしれない。会ったとき。

服部：会ったときにやっぱ対面。やっぱり微妙にコアのメンバーでも、微妙にやっぱり興味の対象っていうのは、微妙だけど、明確にちょっと違いはあるってことですね。

斉藤：全然違うんじゃないかな。

服部：何かみんなが一緒っていうんじゃなくて、何かそれぞれ結構、ばらけていてみたいな、何かそんな。

斉藤：だから、面白かったんだと思う。言語のことなんか、私、その場で勉強したことのほうが多いんかもしれん。

大澤：この辺でもやっぱり橋田さんとか、白井さんはすごく言語っぽい話を。

斉藤：そうだよ。

大澤：そう。聞いてきた論文でも多いなという印象で。

斉藤：そうそうそう、そうそうそう。

神崎：逆に伺いますと、そんな、割と皆さん、多分、認知だとか、言語だとか、会社、異なっていた人たちがなぜ人工知能の研究会で集まられる感じになったんでしょうか。

斉藤：何でだろうね。

神崎：例えばこれが認知科学研究グループとかで集まっても良かったということですか。

斉藤：単に AI。単に私が AI かぶれたって言ったから、かぶれた AI の名前になっちゃっただけかもしれない。認知科学にかぶれたって帰ってきて言ったら、認知科学何とか。でも、エセックスには認知科学はなかったかもしれない。いや、言語学の人がいたんだよな。言語学の人はいたけど、認知科学はいなかった。

大澤：ただ、これだけの広い人が集まれる枠として、人工知能ってのは結構、キーワードな。

服部：えたいが知れない(笑)。どういう。いろんな側面が、ちょっと考えると、もういろんな捉え方があるよねっていう、何か広かったからっていうことですかね。

斉藤：そうかもしれないね。私はそういう感じはしていて、初期の、特に初期の AI の論文って、もういろんなテーマが出

てくるんですよ。つまり対象……。何でもつくってみよう、だから、AM みたいな数学の概念をつくります、みたいな論文があったり、何か絵を描きます、みたいなのがあったり、何かこう、分野的にはもういろんな分野の論文が出てくるじゃん。だから、言語の話もあるし、音声の話もあるし、だから、何ていうのかな。そういう意味で、もともと AI の分野がバラエティーに富んでたからかもしれない。

服部：結局、最終的な対象ってのは人間で、もう間違いなくいると。なんだけど、何かすごくよく分からない。捉え方もいろいろあってっていうところで、やっぱり集まりやすかったってこと。これが何か言語とかっていわれると、ここまでバラエティーをもって集まるってこと、もしかしたらなかったかもしれない。

斉藤：そうなのかな。分かんない（笑）。

大澤：でも、プログラミングの。すごくコンピューターアーキテクチャーみたいなところから。

斉藤：そう。だから、プログラミング……。集まった人の中には、全然、AI じゃない人もいたんだよ。全然っていうか、秋山君なんて、私の知り合いだったから、入ってきたんだけど、精神科のお医者さんだし。

大澤：そうですね。サイコロジーの論文がすごく。

斉藤：あと、來住さん……。そういう意味では、松方君もそうかな。松方君、AI っていうよりは計算機屋さんだったから、持ってくる論文も何か、そういう論文だったけど、別にそれは全然お構いなく。

大澤：なるほど、なるほど。じゃ、みんな、面白がれたっていう感じなんですね、多分。

斉藤：そうそうそうそう。

大澤：そこがすごい。

斉藤：持ってきたものをみんなでああだこうだ言って。これ、何だっけ。6 時までやるんだっけ。

江間：6 時には。でも、お店はもう予約されてるので、6 時半。

服部：6 時半で。

斉藤：もし良かったらですが、テープのね[56]。つまり、実際の AIUEO、どうだったかを聞けるので、良かったら、それを聞きながら、ちょっと喉でも渇いたら、それ、みんな用にと思って、買ってきたので、飲みながら話を続けませんかと。これが何だっけ。音の質が……。始まる。

斉藤（録音）：前もって題を言っとかなかったんですが、そこに書いてあるように、Is it possible to represent symbolizer symbolically?という題で話を。

斉藤：これ、私の声。

江間：先生。

大澤：若いですね。

斉藤（録音）：それで、いつもと違ってペーパーはないんですが、あまり私が話してるよりも、みんなで議論したい。

江間：お幾つ？

---

[56] 1984 年の AIUEO のテープは別途公開予定

斉藤：31 かな。

斉藤（録音）：前半、どういうことを議論したいかという話をやって、で、後半 1 時間ぐらい、みんなに議論していただきたいと思います。それで、だから、話の目的は大体 3 つぐらいあって、最近、私、頭ん中でもやもやしてて、自分で何だかよく分かんなくなってきてることをみんなに議論してもらいたいのがまず第 1 です。それから、AIUEO、今までやってきたんですが、人のペーパー、読んでるのもそろそろつまんなくなってきたから、自分で問題を提出して、考えるような方向にこれからだんだん行くといいのではないかと思うので、その example を提供しようと思って、ペーパーなしでやることにしました。それから、レジュメの中に少し含まれてますけど、通研で最近、リンケージというのがはやりだして、つまり部屋の枠を超えて、他の部屋の人と集まっては議論するっていうのなんですけど、その場でいろいろ議論したことがネタになってますんで、通研で、今までの研究とちょっと違って、そういうばかなこともやりだしたと。民営化すると、…。そこら辺が話、できればいいと思います。それで……。

○○[57]（録音）：もしかして斉藤さんが行ってる間に民営化されたりして。

斉藤（録音）：そうね。

斉藤：彼、誰だか分かんないんだけど、行ってる間っていったら、私、フランスに行く。この直後にもうフランス行くことになってたんだと。

斉藤（録音）：それで、最初のページに AIUEO って書いてあるやつです。

斉藤：この資料は皆さんにお渡ししたよね。だから、見てもらうと分かるんだけど。さっきから話してたように、シンボライザーの話を。そこが問題なんだってことを一生懸命、議論しだしたんですけど。

斉藤（録音）：他の 2 つは参考資料で。最初のやつも 2 ページだけがあれで、その下、参考資料なんですけど。「状況設定」っていうふうに書いてあります。だから、今日は議論したいんで、だから、私がなぜ頭がこんがらがってるんだっていうのを説明して、みんなに解き明かしてもらおうと、そういう感じです。まず 1 番目に、「History of AIUEO」って書いてあって、それは資料の 1 っていうのがありますので、ちょっと見てください。これ、AIUEO から今日まで、最初のほうはちょっと記憶が定かでなくて。っていうか……。

斉藤：ここら辺はずっと、みんなにもお渡しした資料を、私が説明してるだけなんで、ちょっと面白そうなところを。ちょっと飛ばしてね。ここら辺から。

斉藤（録音）：＜聞き取り不能＞

江間：黒板に何か書かれてるんですか。

斉藤：何を書いたかっていうと。

服部：覚えてる、覚えてる。

斉藤：覚えてるっていうか、これを聞いたから思いだしたんだけど。何を書いたか。

---

[57] 以下、話者が特定できない発言は○○で表す。

その頃ホフシュタッターの本とか、まだ出てなかったのかな。一生懸命、アナロジーとかって言ってたじゃん？通研でもそういうのがはやったりして、「次の文字はなあに」っていうのを。こうでしょ。

＜ホワイトボードに以下を書く＞
［ＡＦＨＩＪＫＬＭＮ ○ ］

斉藤：いや、邪魔だ。よいしょ。かな。「次は何でしょう」っていう問題を、シンボライザーの話に入ってるんだけど、出したんですよね。これ、なかなか難しいから[58]。

江間：Ｒ？

斉藤：Ｒ？　違うな。

江間：違う。

斉藤：何でＲ？

江間：何か……。でも、違うな。それ、Ａがそうしたら入らへん。何か、音的な話なのかなと。

斉藤：え？

江間：音的な話なのかな。

斉藤：音？

大澤：Ｅは。Ｅは入ってる。

斉藤：まあ、いいや。これ、考えだすと、多分、あれなんだけど。ちょっとそこの、これ、書いた後に言っていて。

斉藤（録音）：そのときに、人間は文字を認識して、その結果、これを考えてるわけね。だけど、これはこの問題は、番号読み取り機では絶対解けない。

○○（録音）：（　）シンボライザーじゃない？。

斉藤（録音）：じゃなくない。だけどつまり人間はこれを読み取ったときに、もちろん何て言うの、アルファベットだったら、アルファベットは26個あるうちのどれであるかを区別しなくちゃいけないと。そういう意味でのシンボライズをしてるわけ、当然。この中のどれであるかを判定することが、それは単に１つの仕事だけであって、そうじゃなくて、これを見た途端に、人間はこれの発音とか、それから形に関する情報も何らかの形で頭の中に出来上がってるわけで、これを考えるときに全部使えるわけ。そういう情報を。そういう意味で、何て言うか、私はシンボライズという言葉を使ってて、だから「認識」と言ったほうがいい、「シンボライズ」というのは言い方が悪いのかも。

平賀（録音）：違いますよ。だけどそれを見た途端に純粋にＡというシンボルを取り出す過程があって、そのＡに関する知識というのは別のところにあって、それはＡというものが認識され、認識した段階で、それと結び付けられて、そっからいろんなそういうＡに関する情報が出てくるってことじゃないですか。

斉藤（録音）：なるほど。そう思う？

斉藤：これ、平賀君。

斉藤（録音）：知識というのが、何か別の場

[58] 答えはＰ。１つ１つのアルファベットを両方の手で上下から挟み込むようにすると「手が痛い！」アルファベットとのこと（上か下が尖っている）。

所にためてあって、それへのインデックスが A とか、B とか、C とかなってて、つまりそれでインデックスできると。例えば「犬」っていう。

平賀（録音）:「犬」というのはもっと別かもしれません。

斉藤（録音）：ちょっと複雑になってきたけど、例えば「犬」とか、「おばあちゃん」とか、そういうのを認識するわけ、ぱっと見たとき、どっから来てるのか。そのときにそういう例えばおばあちゃんに関する知識が全部ここに（　）に書いてあって、そのおばあちゃん…。

平賀（録音）：そうは言ってないでしょう。その A でインデックスされるものが頭に（　）。

斉藤（録音）：ただ、こうなってるかもしれないと、こう（　）。

斉藤：ちょっと分かんないかな。画面がないと分かんないね、ここら辺ね。でも、こういう話をして、この 1 つ、私の記憶によれば、私の記憶によれば、この一つ一つのシンボルだよね、もうね。明らかに頭ん中にもあるよね。でも、そのシンボルに付いてる周辺情報って、すごいよね。この問題を解くために、何かとっても大変なんだけど、後の……。

大澤：そういう感じですか、みたいな（笑）。

斉藤：後の楽しみにしときましょう。それで、そういう議論をしたり、あと、これが 1 時間。もう 1 時間過ぎたあたりだから、議論に入ってるところで、音楽の話が出てくんだよな、これ。これも平賀君だったかな。

斉藤（録音）：精緻にできてて、人間は。それのお陰で世の中認識できてるわけ。認識した後、後っていうか、頭の中で考えたり、推論したり、三段論法なんていうのは、大体 1 段か 2 段しか人間できないんだから、そこはもう Prolog ですぐできる話なわけ。そんなことは Prolog にやらしとけばいいんだけど、ここは書けないわけ、なぜか。難しいわけ。

斉藤：三角形の、さっきの下のほうのこと言ってる。

○○（録音）：おばあさんというシンボルを作ったまでゴチャゴチャっとあるこの途中のはシンボルにはしない？。

○○：記述できない。

斉藤：うん。記述できない。プログラミングという手段が……。

斉藤（録音）：だから途中からこうなってるのもあるかもしれないわけ、よく分かんないんだけど。でもとにかくこういうプロセスが計算機には欠けてるし、書こうと思ったときにどう書いたらいいのか、私には分からない。

斉藤：今でも分からない。

斉藤（録音）：というか、もうここ 1 カ月ぐらいおばあさんシンボライザーを書こうと思ってんだけど、何をしたらいいのか全然分かんない。

○○（録音）：（　）。

○○（録音）：言語のシンボルはシンボルとしてのものですか？

斉藤（録音）：もちろん。だから私は 5 つぐらいあるかと思うんだけど、さっき言ったベーシックカテゴリーがあるでしょう。それから文字。文字が何千種類あるか知らないけど、それぞれにあ

るわけ。あとワードにも多分あるわけ。だけどそのワードのレコグナイザーは、おばあさんというワードは、おばあさん、字面で見ておばあさんと思うでしょう。だから現実のおばあさん見たときのレコグライザーと非常につながってるわけ。でもその各々に全部あると。そうしないとっていうか、そうすると、その後もう、ほとんどのことはできるよね。という気がするんだ、私。

平賀（録音）：私はこの前ちょうどその話をして、音楽というのは、楽譜レベルというちゃんと立派なシンボライズされたものがあるから、とりあえずはそこを出発点として、話は始めようではないかということをお話ししてきたんですが。

斉藤（録音）：うん。

斉藤：これ、平賀君。平賀君、音楽情報処理みたい。

斉藤（録音）：何？人間が取り扱う対象でということ？

平賀（録音）：そうそう。

斉藤（録音）：取り扱ってる途端にもうシンボライズして取り扱ってんだ、人間は。ほとんどのもの。。

平賀（録音）：意識的に取り扱う。

斉藤（録音）：意識的。無意識的にシンボライズされてるっていうの？　音楽は？

平賀（録音）：音楽がシンボライズされてるかっていうか。

斉藤（録音）：音楽自身が？　人間の中で？

平賀（録音）：うん。

斉藤（録音）：うん。されてるんじゃない？

平賀（録音）：それはされてると思いますけど。

斉藤（録音）：例えばメロディーっていうのがここに来てもいいわけ。

平賀（録音）：ただ、言語の場合には、この上に来るシンボルのその区切りみたいなのかなりはっきりしてて、単語とか、フレーズとか切れるけども、音楽の場合にはそれは切れない。かなりあやふやになってくるから、そういうシンボル化の対象としては、もうちょっとシンボライズするレベルで問題が出てくるわけですね。だから言語を対象にした場合には、言葉というのが切り出せれば、一応そこでシンボル化できたなと思えるけども、音楽の場合にはどっからどこまでを切ればこれが 1 つの単位として認識できたかということ自体はあんまりはっきりしない。

斉藤：まあまあ、こんなことを議論してたわけですね。

大澤：いや、いいですね。

斉藤：え、いいですか。私ね、一番面白かったのは、この中川君の発言なんですよね。これ、私は今でも。

大澤：貴重な資料（笑）。

斉藤：貴重な資料。

大澤：いや、本当に。何ていうんでしょうね。

斉藤：これ、そのうち、文字に起こして、何かどっかへアップしたいなって思うけど。

大澤：いや、もう音声のままでも学生に聞かせたい。

斉藤：音声のまま。

斉藤（録音）：そんな気するね、確かに。

斉藤：中川君。

江間：中川先生、理研のですか。
斉藤：中川先生。
江間：若かりし声。
斉藤（録音）：それは非常に少ないので、作っても売れないわけ。ナレッジエンジニアリング、一生懸命売ってるけど、売れないんだもん。
平賀（録音）：売れないのはいいけど、ただ少ないっていうのは？
○○（録音）：だからほとんどの部分はこっちでやってる部分という。
＜発言が重なり聞き取り不能＞
○○（録音）：だから本当に理論的な問題にアプライしようとすると、全然アプライできないのは、ここをみんな人間は使ってやってるから
＜発言が重なり聞き取り不能＞
相田：エキスパートの知識を一生懸命ピックアップしてっていうのは（　）。
中川（録音）：人間が今の脳の重さを獲得してから概念を獲得する能力を得るまでにかかった時間に比べて、言語を習得してコミュニケーションをできるようになってからの時間のほうがはるかに短いわけだから、それとの対比で考えて、その点は正しいと思う。
斉藤（録音）：なるほど。なるほど。
＜発言が重なり聞き取り不能＞
中川（録音）：正しい、非常に難しいと思うんですが。
＜発言が重なり聞き取り不能＞
斉藤（録音）：そうだね、それは言えてるよね。明らかにそうだ。進化っていうのをちゃんと時間のスケールを押さえとかないと、恐ろしい。
中川（録音）：やっぱり難しいことをやると時間かかるんだと思えば。
＜発言が重なり聞き取り不能＞
斉藤：ちょっと遠くの人の声がちゃんと拾えてないんだけど、でも、こんな議論をして。私、あの中川君の話、聞いたんで、こっちが多いぞっていうのが正しいなと、そのとき、思った。全体の処理の中で。
大澤：シンボル化する過程なんですね。
斉藤：シンボル化して、後で、中でシンボルでマニピュレートしてる部分って、そんなに多くなくて、それ以外のところのほうが大きいだろうと。つまり進化の時間でいうと、たくさんの時間をかけてつくり上げられてきた部分。分かんないけど。だって、動物でも状況の認識とかってすごくできるよね。犬は主人の様子見て、何か大体、やるべきことをちゃんと判断できるとかさ。だから、言葉を使って、人間がやってるって……。何ていうのかな。人間は言葉しか意識できないから、ここがすごいことだって思うけど、実はこっちがすごいっていうのが言いたかったことで、それ、シンボライザー仮説とかっていって言った。
大澤：ある種、動物もシンボルを持ってるってこと（　）それは。
斉藤：もちろん。
大澤：なるほど。
斉藤：違う？
大澤：いや、それは……。何をもってシンボルというか。
江間：そうそう。定義の問題。
斉藤：だって、サル山の彼女、誰だっけ。おサルさんはちゃんと数だって認識す

るよな。あれ、すごいスピードで認識するし。数。数っていうシンボルは持ってる。あ、分かんないな。人間のような数じゃないよ。だけど、例えば自分の主人なんていうシンボルは絶対持ってるよね。遠くからも、匂いかもしれないけど。だから、その基になるセンソリーデータが何かは違うけど、ものによって。でも、それを基にして、この人は自分の主人とかっていうの、認識して行動してるよ。違う？　言語ではない。

大澤：いや、そうだと思いますね。ただ、共有されないってことは、もうちょっと広い……。

斉藤：共有されないかな。他の犬と。

大澤：いや、うーん。そうですね。サルはちょっとそういう意味で、シグナルを送り合うから、結構、インテリジェントな。

斉藤：もちろん。

大澤：つまり言語を持ってないと言っていいのかがちょっと。

斉藤：鳥だって、イルカだって。

大澤：言語持ってるという言い方もできて、ちょっと今、悩んでるんですけど。ちょっとお待ちください（笑）。でも、要するにサルにリンゴ見して、別のリンゴ見しても、まあ、リンゴっていって、同じ反応するんであれば、ある種、シグナル化はできてるってことですよね、多分。このリンゴじゃない。別のリンゴとか言わない。

斉藤：ソーシャルな動物はみんな、アリンコだってできてるよ、きっと。アリンコがシンボルを持ってるかどうか、ちょっとまた、全然、議論が分かれると思うけど。

大澤：アリはちょっと。昆虫は結構、自動機械的に扱われることが多いですね。

斉藤：そうだよ。だけど、自動、自動機械で、こういう仕組みを使ってて、全然構わないよね。

大澤：ああ、もちろん。

斉藤：そこから進化してきて、最後、人間になってるから、ベースの仕組みとして、そういう何か、センソリーの入力情報を少し抽象的な情報にまとめ上げるっていう仕組みが脳の中にある。ちっちゃな脳かもしれないけど、アリンコ。

大澤：でも、そういう意味でいうと、そもそもニューラルネットワークそのものがある種、アナログなデータを、ある種、（　）。

斉藤：もちろん。そういうことは……。

大澤：それをシグナル化というなら、シグナル（　）。シンボル化か。

斉藤：シグナル化。シンボル化。

大澤：シンボル化になる気もしますが、まあまあ、そうか。とか言って（笑）。なるほど。

斉藤：まあまあ、そんなぐらいかな。この先、あんまり面白くは。

江間：何人くらいの人？

斉藤：このとき？

○○（録音）：手の動きとか、体の動きっていうのは何百万年も進化してきてる。それに比べて医学なんていうのは、まだ数百年しかないから、だから書けるんだって。

斉藤：でも、十数人かな。

斉藤（録音）：だから書けるね。さすがそう

いう。
斉藤：まあ、延々、こんな感じで。
江間：口調が皆さん、フランクですよね。
斉藤：もちろん。そういう会だったから。
嶋田（録音）：やればもう一触即発っていうか要するにすぐできちゃう。
＜発言が重なり聞き取り不能＞
中川（録音）：だって AM なんてメチャクチャ動かして出しちゃうけど。
平賀（録音）：僕はあれは十分大したものだと思う。
中川（録音）：しかできなかったというか、でも新しいものは見付けていないでしょう、少なくとも。
＜発言が重なり聞き取り不能＞
平賀（録音）：しかしそれは対象が対象だったから。
中川（録音）：もしだからあれが新しいものをめっけるようになったらば問題は非常に面白いんですが。
斉藤：だから、前に読んだものとかがベースになってて、議論ができてるんでね。
平賀（録音）：しかし逆に子どもがああいう状況で放り出されて、新しいもの見付けるかっていったら、それは難しい。
斉藤：だから、この議論を延々やってたので、あんまり。あんまりこう。
大澤：いや、いいんじゃないですか。議論したい、すごく。
神崎：これ、何年ぐらいの？
斉藤：これ、1984 年 7 月 21 日。
大澤：まだ私が 2 歳のときですから（笑）。
久木田：毎回、取ってたんですか、録音。
斉藤：いやいや、違う。これは、だから、私はもうこれで、この会で私、1 年間、フランス行っちゃうから、AIUEO に出るの、これでしばらく終わりだなとかって思って。かつ、自分の考えたことを話そうとしてたから、なぜかカセットを持ってって、録音してたんです。しかも、そのテープがずっと忘れてて、あることももう忘れてた。あるとき、部屋の整理を、子どもが独立して、何か部屋の整理をしてたときに、昔のカセット箱みたいなとこから出てきて、えっ、こんなのあったんだ。かけてみたら、聞こえたから、取り込んで。
大澤：いいですね。やっぱり何かこう、生々しいし、何かこう（笑）。書き置こしてると、失われるものがあるじゃないですか、結構。
斉藤：いや。だけど、私は逆に、これ、あったんで、今、っていうか、もうそろそろ、私、京大を定年なんだけど、定年なった後、やることが何か、それのきっかけになって。FUN-AI もそうだったんだけど、あのときも昔のことを思い出すのに使えて。で、思い出したし、今も何か、ちょっと皆さん、来るからっていって、昔のもの、読みだしたら、え、こんなこと、書いてたんだ、みたいなので。もうどんどん忘れてるね、人間はね。いや、すごい量を忘れてるっていうか、覚えてるのが少ないんだよね、多分ね。
大澤：いや、そうですね。割とそういうことある気がします。
斉藤：なので、それがよみがえって、案外、面白い。どうぞ。
神崎：今、聞かせていただいた部分だと、ものすごく行儀の良い会に聞こえる。
斉藤：行儀の良い。

神崎：つまり発話のオーバーラップはあるんですけれど、それによって、誰かの発言がキャンセルされたりっていう部分がない。

斉藤：キャンセル？

神崎：つまり何か言おうとしたとき、誰かかぶせてきて、こっちの人が言おうとしたのができなくなるとかっていう。

斉藤：そんな傲慢（ごうまん）なやつはいなかった（笑）。

神崎：そういう雰囲気だったですね。いや、つまりその研究会の雰囲気として。

江間：聞こうとする姿勢がある。

神崎：そうそう。つまりものすごく発話者と次の発話者の切り替えがスムーズな会合。

斉藤：へえ。

江間：あと、話題が多いなっていうのは。何か和やかなというか、フランクに進んでいるのかな。ぎすぎすは全然してない。

斉藤：ぎすぎす。

大澤：みんな、聞きたいんだなっていうふうに思う感じですよね。みんながみんなの意見を。

斉藤：みんな、ぎすぎすする会があるの？

大澤：私も気になったんです。そんなとこ、あります？（笑）

服部：結構あるね。

斉藤：結構ある。

江間：でも、こういうの見てると、最初は結構、月曜日とか、金曜日とか、先生のまとめ資料の時間も書いてあるので、だけど、もう後半は土曜日の3時5時でほぼ固定されてるわけですよね。

斉藤：そうだね。だから、それは何か、就職した人とか出てきてからでもないかな。大体、土曜日に落ち着いたんだよね。

江間：はい。で、何か、じゃ、この後、飲みとか行かれてたんですかっていったら、そんな感じでもないみたいなことを、何か誰かおっしゃってたような気も。

斉藤：飲み会……。飲み会はどうだったかな。あんま覚えてないな。

江間：先生のこの資料、資料1って書かれてるんですけど、これは、この最後、もうフランス行かれる前に、今までのまとめを出しますよって意味での資料1なんですか。

斉藤：そうです。これ、今のテープの話のときに配った一連の資料の一部。

江間：はい。何でこれを出されたんですか、この資料を。毎回、配ってたわけじゃないですよね？

斉藤：違う違う。最初に言ったみたいに。あ、これか。これがそのとき、実物なんだけど、実物って、これ……。

江間：はい、はい。あ、手書きだって感動したのをちょっと、今日。そうだよなって。Word なかったよなって。

斉藤：そうなんだよ。まだ TeX もなしみたいな感じで。

江間：先生の字、きれいだなっていう（笑）。

斉藤：この状況設定っていうところで、今までこういうこと、やってきましたっていうのを言うために。それと、忘れちゃうのもあって、良くないなとかとも思ったから、これは打ち込んで、それを持ってったの。だから、ここの最初で話すために、この資料は使ったの。

江間：なるほど。それで、資料 1 参照って、そういう意味で。なるほど。

斉藤：うん。それで、何でシンボライザーっていうの、大事だと思ったか、みたいな話をして、この、何だっけな。T4 っていうバクテリオファージの話がとっても面白いんだけど。

大澤：何かそれ。それもかなり複雑な。

斉藤：いや、それはその本を、本を読んで、面白かったから、ここで紹介したんだけど、この IJCAI の……。でも、やっぱりマー[59]の、マーの論文が、こないだ、週末に読み返したら、マーの論文が面白かったな、やっぱりな。デビッド・マー、あれが最後の遺言みたいな感じなんだよね、白血病で死んじゃったから。本当に……。で、私は彼の言ってる、案外、ボトムアップに下から積み上げて、分かるところから分かってっていうふうに共感できて、そういうふうにやらないと嫌だな。そうじゃないペーパーも多かったんでね、AI のペーパーってね。シャンクもそうだし、ミンスキーもそうだし。ミンスキーはどちらかっていうと、やっぱりアーキテクトだよね。そう思わない？ フレーム理論っていって、「こういうフレームを持ってるから、人間は知的なんだ」なんて言っちゃうわけだから。えっ？　それで全部、（　）できないよなとかって思うし。そうなんです。だから、マーが言ってたものも関係してるし、バクテリオファージの話は難し過ぎるぞって。人間の脳でやってることって、もっと難しいぞって。こんな簡単な生物でも、こんなに複雑なんだからっていう意味で話したんだけど、これ。これ、面白いよね。今、どこまで分かってるのか、知らないけど。こんなに小さいのが遺伝子で動いてさ。しかも、大腸菌の中、入り込んで、大腸菌、全部、食いつぶして、自分、どんどん増えていく、とかっていう話をして。それを焼き直しただけなんですよ、FUN-AI で、話したことは。大昔のこれをひっくり返して、昔のテープも見つかって、聞いて。そうだ、こんな話してたんだ。あのときは、とかって思って。ほんで、話したら……。だから、中島も……。松原君、ここにいるんですよ。松原君、出てくるから、言うけど、何かあんまり覚えてなかった。あ、ごめん。中島はいないんだ、このとき。このときは、中島は MIT 行っちゃってたんだよ。

大澤：ああ、なるほど、なるほど。

斉藤：1 回目の MIT か何かに行っちゃってて、中島、いないの。だけど、松原君もいたし、橋田君もいたし、みんな、案外いたんですよ。

江間：だから、その資料が、資料 1 がつくれるっていうことは、今までつくりためてたっていうより、この発表されるときに改めて打ち直したみたいな？

斉藤：これは、私、メールを、さっきのメール残ってた……。あ、違うな。こんな最初の頃は、そうだな。あ、どうしてだろう。ごめんなさい。どうしてこれ

[59] David Marr. イギリスの神経科学者。計算論的視覚論で知られる。

が残ってるか、よく分かんないな。メールをやりだしたのは、もっと後だもんな。何かメモってたのかな。あるいは、案外、物持ちがいいので、そのとき、配ってもらった紙を。

江間：資料、全部、見返して。

斉藤：取っといたのをめくりながら書いてったか、どっちかだな。

江間：結構、他、中島先生とか、堀先生とかも何かお話、聞いたときに「次、どなたにこういうお話、聞いたらいいでしょうかね」って言ったときに、やっぱり斉藤先生のお名前が挙がる。どこでも挙がったんですけれども、それは皆さんもやっぱり斉藤先生がいろんな資料を持っているよっていうふうには思って。

斉藤：ああ。FUN-AI のときにだいぶ分かったんじゃないのかな、それが。それまでは知らなかったんじゃないかな。一応、言いだしっぺなので。

江間：幹事っぽい役割だった？

斉藤：やってたと思います、最初の頃は。途中で松原君に渡したのは、それこそフランス行く頃かな。どっちかな。もっと後。

江間：でも、先生が今、持っていらっしゃる資料っていうのは、今、今回、事前に頂いたこの辺の資料っていう。

斉藤：そう。最初のほうだよね。だから、この後、84 年から 91 年ぐらいまでかな。5〜6 年、まだ続いてるんだけど、そこはだんだん、私も何か忙しくなって。合宿とかは行ってたけど、毎週のやつはあんまり行ってなかったかもしれない。だから、後半はちょっと分かんない。

江間：なるほど。でも、この前半部分のいろんな資料とか、何かまともに多分、本当に読んだものとか取ってたら、段ボール、結構な数になると思うんですけど、そういうのはもう、いろいろ移られる中でどこかにとか、そういう形ですか。

斉藤：うん。どっかへ行っちゃってるよな。引っ越したんですよ。武蔵野にいて、厚木に引っ越した。それが 1 つのタイミングで。厚木に引っ越したのがいつかな。厚木に引っ越したのがね……。ここか。92 年か。でも、だから……。だから、そのときに、92 年に厚木に引っ越したときに、だいぶ資料としては捨てちゃった。

江間：なるほど、なるほど。

斉藤：厚木に残ってたのは、ここに持ってきてあるんだけど、この後ろにあるんだけど、それは囲碁の研究とかしたときの資料で、紙はあるんだけど。あんまり昔のは……。でも、昔の少し残ってて、これなんか、そうなんだよね。文系の人との、文系の人とのやりとりっていう意味では、私はこの情報文化研究フォーラムっていうところで、松岡正剛とか、そういう人たちとやりとりしたほうが、何か文系の人と。

江間：先ほど、お話された。

斉藤：寺子屋の話、出てくるじゃない？　江面[60]先生、江面さんの。それじゃなく

[60] 江面竹彦。産業図書社長。2003 年逝去。

て、これの、ここで、安西先生なんかも出てるんだけど、黒崎先生とかに会ったのは、私、ここだと思うの。

江間：へえ。

斉藤：何ていうのかな。文系の人との交わり。荒俣宏[61]なんか、すげえおじさんだったな、この人。

大澤：そうそう。ちょっとびっくりしました。

斉藤：え？　荒俣宏、この人、すごい人だよな。

江間：っていうか、西垣先生[62]がコーディネーターなんですね。西垣先生。

斉藤：そうそうそう。これ、面白かったですけど、かといって、AI の話すると、何かけちょんけちょんに言われて。それは誰だった……。

大澤：ええ？（笑）　AI と言語って書いてある。

江間：これは第 6 回って書いてありますけど、全部で 6 回やられたってことなんですか。

斉藤：いや、もっと何かあったかも。1～2 年……。

江間：一応、第 1 回から？

斉藤：1～2 年、つながってて。でも、そこで終わってるなら、それぐらいかもしれない。2 カ月に 1 回ぐらいかな。

江間：へえ。

大澤：これは年としてはどの辺？

斉藤：年は 2001 年ぐらいじゃないかな。

江間：へえ。

斉藤：文系の人って立て板に水で、頭ん中からどんどん言葉が出てくるんだね。いや、本当に。それはすごくびっくりした。何か歴史の話とか聞いても、年代も含め、人の名前も含め、もうすごい細かい情報が全部、頭に入っててさ。もう立て板に水、どんどん言葉が出てくるのはびっくりしたけど、でも、そこで議論してて……。だから、AI の話、何でけちょんけちょんに言われたのか、全然、覚えてないんだけど。

江間：でも、この会合の名前に AI が付いてるんですよね。

大澤：そう。AI って付いてるのに、けちょんけちょんにされる。

斉藤：そうです。けちょんけちょんっていうのはね……。

江間：へえ。AI と言語。誰がこれ、お題。松岡正剛さんが。

斉藤：いや、NTT の、遠藤力さんっていう人が NTT の本社の何か企画部みたいなところにいて、そういうの、企画したんだよね。企画して、松岡正剛に頼んだ。

江間：はい。

斉藤：松岡正剛は文系の人たちを呼んでくる。AI だけじゃなくて、何か他にも幾つかセッションがあったらしいんだけど、AI のところに私と竹内さんと呼ばれて。

江間：これ、ゴールは何だったんですか。

斉藤：何だろうな。情報文化フォーラムだから。でも、そこのときので、何か松岡正剛、本、1 つ出してないかな。

神崎：うん。アマゾンにありますね。

斉藤：情報……。

---

[61] 荒俣宏。博物学者、作家。

[62] 西垣通。東京大学大学院情報学環名誉教授

神崎：『情報と文化―多様性・同時性・選択性』という。

斉藤：ああ、何か。でも、それが目的ではないと思うんだ。NTT としては、何の役に立ったんだろうな。何の役にも立ってないかな(笑)。

大澤：ぱっと見ると、かなり人文寄りな。

斉藤：そうだよ。

江間：でも、その『情報と文化―多様性・同時性・選択性』[63]って、1986 年。

斉藤：じゃ、違うよ、それ。

神崎：あと、なんか 1992 年に……。

斉藤：もっと後に、何か情報文化……。

神崎：『問題集』、これですね。

斉藤：それそれそれ。

神崎：これに AI が。目次の中に AI。

斉藤：『情報文化問題集』[64]っていう、何か本を出したのが、ネタは、ネタはその会合だと思う。

大澤：これからの AI が。

服部：メディアとコミュニケーション。

斉藤：何かそういう議論はよくしてたんだよね、その頃。2001 年って、何、冬の時代？

大澤：そうですね。この頃、AI やるっていうのはなんか。

斉藤：そうなんだ。

大澤：ちょっと難しかった。

斉藤：そう。私、そういう感覚、全然なかった。

服部：大学で、僕、「君ら、貧乏くじ、引いたんだよ」って言われて（笑）。もうはっきり言われました。

斉藤：へえ。

＜発言が重なり聞き取り不能＞

斉藤：すごいこと言われてるような気がする。そうなんだ。全然、ちょっと私は、そういう感覚を一度も持ったことがなかった。

江間：ごめん、神崎さん、さっき言ったのっていうのは、タイトルは？

神崎：『情報文化問題集』。

服部：92 年。

神崎：NTT 出版。

江間：へえ。なるほど。本当だ。そうか。

斉藤：まあ、何でやってたんだろうな（笑）。で、何となく面白がって出てったけど。

大澤：デビット・マー（David Ma）の、ここでも引用されてる。

斉藤：そうそう。私、その頃、ずっとマーとか、ウィノグラードも好きだったし、マーも好きだった。ウィノグラードは面白かった。彼の書く英語はとっても分かりやすいんだよね、ウィノグラードの。これがウィノグラードのドクター論文なんだけど、これ、すごいんだよね。

大澤：え、本人？

斉藤：これ、向こうから取り寄せたやつで。昔は、本屋で売ってるようなもんでもないし、彼は後で本にしたけど、『Understanding』何とかって題の本[65]にしたけど、これはドクター論文そのものを取り寄せたやつで、私が入った研究室の、竹内さんが最初に持ってて、私が入る前にもう読んでたんだよ

---

[63] 情報文化研究フォーラム（編）, 1986, 『情報と文化―多様性・同時性・選択性』, NTT 出版

[64] 情報文化研究フォーラム（編）, 1992, 『情報文化問題集』, NTT 出版

[65] Terry Winograd, Understanding Natural Language, Academic Press, Inc. 1972

ね。これ、71年だから。譲り受けて、私が2番目の所有者。誰かもっとAIやるってやつが来たら、渡してあげようと思ってたけど、出てこない、みたいな。「マル秘」なんて書いてある(笑)。これは竹内さんのいたずらじゃない？これも読んだんだよな。ウィノグラードの英語はよく分かりやすい。Normanとかもそうだけど、英語読んでて気持ちいいって感じの文章書いてくれるよね。

大澤：ディスカッションの面白い。

斉藤：みんな、書くのうまいよな、何か向こうの人はね。

大澤：そうですね。文章家ですよね。

斉藤：しかも、たくさん、うわーっと書いちゃうし、すごいよなと思う。思うけど、われわれ、とてもあんなに書けないよな。

江間：何か、いろいろFUN-AIとかもあって、資料、発掘されて、いろいろ思い出されてて、面白いっていう話で、先生、定年された後とか、その辺の、この録音含めて、何か改めて掘り返すみたいなことを、今後。

斉藤：いや、私は囲碁の、囲碁の強いプログラムをつくるのを老後の楽しみにしようとしてたんだけど、強いプログラムがつくられてしまったので。

江間：別の意味の強さですよね、先生の。

斉藤：いや、で、今、考えて。本当にjust an ideaなんだけど、そこにラズパイが……。ラズパイ、みんな、知ってるよね？

一同：はい。

斉藤：ラズパイ、ここでポケゼミっていうのをやってて、学生10人ぐらい、1年生に、ラズパイ使って、計算機のこと教え、インターネットのこと教え、さらにセキュリティのこと教えてるんだけど、これでAlphaGoに勝ってやりたい。

大澤：おお。というのは、何でしょう。つまり計算機パワーに頼らずみたいな話ですか。

斉藤：いや、私は人間の脳の計算機パワーはあれより低いと思っている。

大澤：なるほど（笑）。

斉藤：ここだよ、みたいな。分かんないけど。これはすごいのかもしれない。もうちょっと、それこそ、ここをつくるときには何か、ディープラーニング使って、一生懸命、頑張ってつくらないといけないかもしれないけど、でも、ここは全然、もうラズパイでできるぐらいの処理しかしてないような気がしていて。だから、この下支えがあれば、あいつ一つでAlphaGoよりも強いプログラム書けるかもしれないって。すごい、夢物語なんだけど。

大澤：そこはあれですか。マルチモーダルなっていうか、センサリーは含めないで。

斉藤：いや、だって、囲碁だから、センサリーは簡単なんだよ（笑）。

大澤：いや、相手の顔見てとか、それではなくってことですよね？

斉藤：だって、そんなことしないよ、プロは。プロは、まあ、しなくないかな。

大澤：まあ、プロはしないか。いや、分かんない。将棋は昔、そういう駆け引きのゲームだったというふうにちょっと聞いた。

斉藤：いや、みんな、やるけどさ。接待、接

待碁とかっていうのもあるからさ。

大澤：まあまあ、ありますね。

斉藤：やるけど、でも、真剣にやる人。だって、オンラインでやる。最近、みんな、対戦してるから、プログラムと対戦するのもオンラインでやってるから、顔色もくそもないよね。だから、囲碁っていう限定したもので考えれば、そういうセンサー情報はほとんど要らず、でも、盤面を認識する仕組みは何か要るんだよね。人間の場合は4歳ぐらいからやりだして、プロになる間、蓄積した過去の対戦の記憶みたいな。本当にみんな、覚えてんだよね、彼らね。

大澤：間違いなくそうですね。

斉藤：びっくりしちゃうぐらいな記憶だよ。しかも、盤面もこう、ひっくり返したり、白黒変えても分かるんだよ。分かるっていうのは、次、ここまであるんですけど、次、並べてくださいって言うと、さーっと並べる。何であんなことが覚えてられるのか、私はいまだに不思議でしょうがないんだけど。

駒谷：感想戦とかすごいですよね。

斉藤：すごいよね。でも、でも、あれはやっぱり考えてるから、覚えてるんだと思うんだよ。深く考えてるから、対戦してるときに。つまり理由があって、どっかに打つし、相手の手も、理由があって、そこに打ってきてるというふうに捉えてるんでね、盤面を。

駒谷：丸覚えしてるんじゃないという。

斉藤：違うよ。丸覚えはできないよ、絶対あんなの。ほとんどランダムパターンじゃない。囲碁を知らない人にとっては、黒と白の石のランダムパターンだよね。覚えてるはずないよね。

服部：何か筋があって、ストーリーで覚えてる。

斉藤：そうそう。絶対そう。

大澤：確か、そういう実験があったような気がしますけど。要するに囲碁的に意味がある。

斉藤：チェスでそれはあります。

大澤：チェスで。

斉藤：うん。で、プロは要するに普通のチェスの試合に出てくる盤面はすごくよく覚えるんだけど、ランダムにつくった意味のない盤面は全然覚えられない。

服部：面白い。

斉藤：だから、そこに特化して、頭ん中、構造が出来上がってる、ここら辺の。

江間：ものを何か、暗記するときと似てるのかな。私、楽器やるんですけど、どこで間違えたとか、ピアノの人とやるときに、あそこでこけてとか、あそこで間違えたよね、速くなったよねとか、あと歴史とか覚えるときとかも、何かストーリーとかで何か覚えたほうが、ちょっと若干、分かるか分かんないですけど、多分、棋譜並べてくみたいな、結局、時間経過のストーリーみたいなところ。

斉藤：そうですね。時間方向の。

江間：だから、何か、暗記とかも、やろうと思って、言語でやるもんじゃないから、その辺の話との違いなのかな。「どうやって暗記してんの？」ってよく言われるけど、「いや、覚えるよね」みたいな（笑）。

斉藤：でも、繰り返すよね。

江間：うん。

斉藤：だけど、繰り返しの回数がもう圧倒的にニューラルネットとは違うので、絶対同じ仕組みのはずないよねって思う。

江間：じゃ、それを今後。

斉藤：分かりたいけど、どうやれば。どうやれば分かるのか、分かんない。誰か研究してくれて、分かったって言ったら、もうそれだけで満足しちゃうかもしれない（笑）。

江間：何かわれわれも幾つか、何人かの方にこういう。80 年代のブームを知りたいとか、AIUEO が何か知りたいっていうよりは、今、AI がこんだけ盛り上がってて、結局、でも、80 年代って、何が議論されてたんだろうとかいうことって、あんまり振り返られて、実は……。結構、多分、仲間内ではいろいろ話されてて、連続性があると思うんですけど、よその分野から入ってくる人だったりとか、今の話は未来しか、みんな、しゃべらないんですよね。下手したら、すごいシンギュラリティみたいな。でも、そこがニューラルネットワークから出てくるシンギュラリティの話と、今の先生から出てくるような、この未来の話と、また何かちょっと分岐があって、違うようなところもあるような気がしてて、何かそこの辺はちゃんと、メモ、メモというか、残しておいて、共有するっていうことは、ある意味、ブームがもし終わったとしてもやっぱ大事な試みだし、われわれ自身が今、異分野連携でいろいろ活動してるけど、その当時もこういう、何か和気あいあいとした感じで異分野連携してたの、そのコツって一体、あるいは核って何だったんだろう、みたいなところがあって。今さらこういう説明なんですけど。そう考えたときに、やっぱり AIUEO って一つ、面白いコミュニティーというか、活動だったんじゃないかなと思っていて、「それはまだ続いてるよ」とか中島先生とかには言われたんですけれど、「Facebook で AIUEO があるよ」とか何か。

大澤：はいはい。

斉藤：ああ。まあ、はい、はい。

江間：そう考えたときに、何か斉藤先生から見て、AIUEO って一体、何だったんだろうとか、あるいはもしこういうことにもっと知りたいって思うんだったら、こういう人とか、こういう分野とか、何かこういう研究とか見てたら面白いんじゃない？っていう、何かちゃんと過去を記録というか、継承していくために必要なこと。あるいは継承なんかしなくていいよっていう話もあるかもしれないんですけど。

斉藤：うーん。何か人工知能っていう分野に限ると、何か今は特に、何かマスコミに踊らされてるような雰囲気、強いよね。マスコミなり、何だろうな。だから、第 3 次ブームは明らかに応用。社会に役立つものがつくれます、から出てきたから、特にそうなのかもしれないけど、それまでの知識工学のとき、ICOT はどうだったか、分かんないけど、それまでの AI はもっと何か地道に、本当に知りたいことを知るために

研究してるっていう感じで進んできたと思うので、その蓄積とか、そのとき、何、考えてたかっていうのが残ってないのはとっても良くないと思う。だから、それは今の人も分かった上で今のブームを捉えたほうが良くて、そういうふうに捉えると、何かできる分野とできない分野があるのに、できる分野だけが拡大して、宣伝されて、すごいぞ。シンギュラリティみたいなので行っちゃうぞ、みたいな議論になってるから、そうじゃなくって、もっと根本的に難しいものがいっぱいあって、それは全然、解けてないぞとかっていうのは分かっててほしいなと思う。でも、それは昔の論文読んだり、何かすれば、多分、誰でも分かる、分かるんじゃないかな。

大澤：でも、何かお話し聞いて、やっぱ議論していた過程とかってのはなかなか、論文だけでは難しい。

斉藤：まあ、確かに。

江間：それこそ息遣いみたいなとか、どういう雰囲気だったのかとか。

大澤：これをどう、みんなが見ていたかっていうとこまではなかなか見えないですよね。

江間：それこそ、言葉にならない部分だけど（笑）。

斉藤：でも、でも、AIUEO のコミュニティーもそんなにでかいコミュニティーじゃないから、それで日本の中でも、AI も違う捉え方をしてる人もいっぱいいるだろうから、だから、ごく一部の、もしかしたら偏ってる考え方かもしれないですけど、でも、本質的な問題がこれなんだっていうのは、あんまり反対する人はいないような気がして。でも、それができないときにどう動くかは、人によって違うわけよね。いや、だから、できないから、できることを頑張ろうって言って、応用のほうにどんどん走る人もいるし、何か私なんか、どちらかというと、他の分野ならうまくいくかもしれないと思って、認知科学行ったり、何か脳科学、一生懸命、人の話、聞いてみたりってやるんだけど、それでも何か、まだ駄目だなとかって感じるとかっていうふうに。そうすると、アウトプット、0 なんだよね（笑）。やっぱり何か、分かんないじゃんっていう状態のまま、延々、何か悶々（もんもん）としてるような人もいるし、どんどん転戦して、新しい分野でどんどん活躍してる人もいる、それぞれだよね。だから、どんなコミュニティーでも、同じようなことやれば、誰でもできると思うし、コツったって、別に（笑）。でも、だから、ゆとりっていうか、時間のゆとりは大事だよね。

大澤：そうですねえ（笑）。

斉藤：いや、つまり、これもそうだけど、学生の頃は案外、頻繁に開けたり、夏の合宿ったって、ああ、いいなとかって、4 泊。

江間：え、そんなに？（笑）

斉藤：4 泊したことはなかったかな。でも、だって、厚い本を読んだから。いや、4 泊はしてなかったな。でも、1 週間ぐらい使ったかもしれない。

服部：その合宿で 1 冊読み切るみたいな？

斉藤：そうだよ。これにも出てくるよね。何だっけ。最初に読んだの、マーガレット・ボーデン[66]の『Artificial Intelligence and Natural Man』[67]とか、その次のときは、これ、スローマン[68]の読んだんだな。『Computer Revolution in Philosophy』[69]なんて。何、読んだか、全然覚えてないな。中身、全然覚えてない、本当に。あ、ボーデンのは少し覚えてるけど、あとは、何、読んだ。カール・ポッパー[70]、あとはポッパー、エックルズ[71]、読んだんだ。あと、ほら、何だっけ。ペンローズ[72]の、ペンローズの本とか。

大澤：それは英語ですか。

斉藤：そうそうそうそう。『何とか's New Mind』。

大澤：『The Emperor's New Mind』[73]ですかね。

斉藤：そうそうそうそう。でも、あれ読んで、私はもう絶対違うなと確信した。

大澤：やっぱりみんな、そう思いますね（笑）。

斉藤：つまり、そんな量子とか何だとかじゃなくって、もう私は何となくもっと単純に、人間の頭の中って有限オートマトンだと思ってるから。チューリングマシンなんて、そんな無限なんて入ってるはずないし、そうじゃないです、有限オートマトンですと思ってるから、もう絶対、原理的にはできる。計算機で実現できる。確信してるんだけど、でも、全然できない（笑）。やれば今でもできるし、何か、何か、何ていうんかな。でも、今、何か過去の2回の読んだ私の感想は、ソーシャルな問題とか、倫理の問題とか、何か余計なこと気にし過ぎみたいな感覚です。そんなこと、私の頃はそんなことはこれほども考えてなかった。毛ほども考えてないし、何かセキュリティーと似てるけど、私、インターネットやり始めた頃は、もうセキュリティーって何にもなくて、アメリカの大学でも何でも入れたんだよね。入れたっていうのは、guest っていう ID で、guest がパスワードだったから（笑）。guest、guest っていうと入れるって。で、中、見れて、教えてくれてとかって、そういう世界だったから、今のように、もう何かセキュリティでがちがちになって（笑）。こういう世界が信じられないんだけど。何か、まあまあ、ロボットとか、本当に世の中に出だせば、確かに考えなくちゃいけないことがいっぱいあると思うけど、何か気にし過ぎっていうか、感想を持ちました

---

[66] Margaret Ann Boden. 認知科学研究者。

[67] M. Boden, 1977, *Artificial Intelligence and Natural Man*, MIT Press.

[68] Aaron Sloman. 哲学者。認知科学者。

[69] A. Sloman, 1978, *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind*, Harvester Press.

[70] Sir Karl Raimund Popper. イギリスの哲学者。

[71] John Carew Eccles. オーストラリアの神経生理学者。

[72] Sir Roger Penrose. イギリスの数学者、物理学者。

[73] R. Penrose, 1989, *The emperor's new mind: concerning computers, minds, and the laws of physics*, Oxford University Press.（＝1994, 林一訳『皇帝の新しい心：コンピュータ・心・物理法則』, みすず書房）

（笑）。そんな、どうでもいいって言うと怒られるけど、そうじゃなくて、本質的な、解きたい問題は別にあるんだから、それ、一生懸命、解けばいいじゃんとかって。責任が違う。

江間：気にし過ぎっていうのは、そこまで行ってないよっていう意味なのか。本質じゃ……。今、先生、本質じゃないっていう話をされましたけど、何か。

斉藤：両方。そこが……。

江間：悪用されるまで、技術、そんなに行ってないよって。

斉藤：いや、だって、私がよく言うのは、悪用されるとか言うけど、電源、引っこ抜きゃいいじゃんって思う。計算機でつくってたら、だよ。

駒谷：単純に（笑）。

斉藤：引っこ抜きゃ、動かなくなるよ。それを何か防御する仕組みを中で計算機が考えて、やるみたいな話、SF みたいなのいっぱいあるけど、そんなん……。

大澤：でも、私はその辺は何か面白いからってのはありますね。これが要するに知能的課題に結構、本質的なとこに近い面があるから、倫理の課題が面白いと思ってる面はあります。それは僕の勘ではあるんですが。何でしょう。危ない話とかも結構、出てくるし、あと、止められるったら止められるんですけど、心的に影響与える場合ですよね。っていうのは、割と依存性があるみたいな話とかは割と面白いなって。うそをつく AI もそうなんですが。っていう理由でかかわってますみたいな、私は（笑）。

斉藤：いや、でも、本当に意図みたいなものを持って、計算機が動きだしたら、確かに何でも起こるかもしれない。

大澤：ただ、意図を振る舞うことはできて、そして、完全に意図を振る舞うようにできるっていうことは、それは内面もかなり実装されてないと意図はおかしい。っていうか、そんな人間はわれわれの敵っていうか、特に敵対的な関係になった場合は。と思っているので、敵対的関係というのに面白みがあるというのは確かに。それを、お金もらえる形みたいに、つまり社会的に受け入れられる形でそういう研究的なテーマを結び付ける必要はあるとは思うんですがっていうわけですが、私のこれは方法論。

斉藤：なるほど。

大澤：でも、面白いから。

斉藤：面白いかもしれない。チューリングテスト、もっと高度にして。

大澤：そうですね。チューリングテスト、もっとソフィスティケートな形にするべきだって議論、本当、結構ありますけれども、テストの一つにやっぱりどれだけうそをつけるかっていう。チューリングテストは結構、そういう課題でもあるけれど、もっとスペシフィックにやってもいいんじゃないかなと思うことはありますけど、まあ、その、はい（笑）。

斉藤：戸田正直[74]先生っていう先生がいて、

[74] 戸田正直。1952 年より北海道大学文学部。87 年より中京大学文学部。90 年より中京大学情報科

それこそ野田セミナーとかに来てた先生で、認知科学の分野では大御所なんだけど、もう亡くなられちゃったんだけど、彼は最後の頃、ずっと感情、感情って言ってたんですよ。感情。確かに感情は知能とちょっと違うけど、でも、人間の場合、もう圧倒的に感情的に動くよね（笑）。敵対して怒るとか、そういうのは特にそうだし、だから、感情、感情を埋め込めるとかってか、感情的に反応する計算機とかって面白いかもしれない。

大澤：いわゆる認知科学でいう感情の役割の一つとしては、例えばいわゆる計算機能をそれでスキップできたりとか、ある種の、要するに方略の一つとして持っておくっていうのはあると思うんですけど。

斉藤：方略ね。

大澤：方略。要するに怒ることによって、いろんなセンサリーをストップして、目標に向かうみたいな、ストラテジーがそのまんま、感情っていう形になっていくみたいな。安西先生とか、よくそういう言い方をされますけど。もう一つの、つまり感情を装うことによって利益を得るというような関係というのも恐らくあって（笑）、そっちはもっと。

斉藤：人間の場合はあるけど。

大澤：いや、それが、それが結構大きいような、人間の場合は。

斉藤：いや、でも、私、全然、別の話になっちゃうかもしれないけど、研究所にいて、途中から事業部に、OCN っていうところに移ったんですね。それはインターネットのことをできる人とかっていっても、NTT としては何かそういう人が研究所にいっぱいいたわけですよ、私も含め。そういう人、がーっと事業部行って、OCN つくる仕事に入ったんだけど、そのときにすごくびっくりしたのは、っていうか、そこで私は、何かこう、普通の会社でのマネジメントみたいな、少し勉強になったんだけど、やっぱり装わなくちゃいけないんだよ。装うっていうか、感情を自分で、演じなくちゃいけないんだよ。つまり部下が失敗したときに、へらへらして「まあ、いいよね」とかって言ってちゃ駄目で、本当に怒らなくちゃいけないんだよ。

大澤：そうですね。

斉藤：怒るっていうか、叱らなくちゃいけない。それはなかなかできなかった、最初。研究所にいると、そういうシチュエーションってほとんどなくて。

大澤：ああ、なるほど。

斉藤：つまり勝手にやっとけ、じゃん。何となく基本は、研究って。

大澤：研究はそうなんですよね。学生に向き合うと、またいろいろ難しい問題が（笑）。

斉藤：そうかもしれない。そうかもしれないけど、とにかくそのとき、私は最初、何か、要するに自分、自分を装う。だ

学部。99 年より梅村学園学術顧問。2006 年逝去。中京大学人工知能高等研究所ニュース No.19「戸田正直先生追悼号」参照のこと。http://www.iasai.sist.chukyo-u.ac.jp/pdf/iasai_news19.pdf

から、怒ってる振りをしなくちゃいけない。

大澤：そうですね。

斉藤：本当に怒ってないんだけど、本当に怒っちゃうことは、私は少ないので(笑)。それは何か、とっても難しかったね。

大澤：社会的な機能として、シグナルを出す。

斉藤：そうそうそう。それで、大勢の人がいたりすると、もっとだよね。たくさんの人に向けて、叱る方向のことを言わなくちゃいけないとか。褒めるんなら楽だけどさ。

大澤：そう（笑)。難しいですよね。

斉藤：いや、本当に。それはなかなかね。

大澤：いろいろ間違えると傷を負いかねないとか、いろいろ考えつつ。そう。それって、でも、インテリジェントだなと僕は思うんですが(笑)。そういう。

斉藤：ああ、なるほどね。感情、コントロールできる人はインテリジェントだよね（笑)。

大澤：いろんなインテリジェンスがあっていいと思います。

斉藤：確かに。

久木田：ちなみに僕は、感情って、コントロールできないところが本質だと思って。

大澤：でも、多分、ある程度はできるって、そのせめぎ合いが多分、ある。

久木田：ある程度はできる。つまり、感情って、行動と、それから表出と両方に影響を与えるでしょ。だから、その表現を見て、あ、この人は次に行動するんだなっていうふうに、社会の中で、人がどういうふうに次、行動するかっていうのを予測しやすくして、何か協調とかを促進する仕組みだというふうに思うんですよね。

大澤：うん。だから、それは半分かなという。つまり、ただ、結局、そういうふうに、みんながシグナルを共有した時点で、裏切ることに合理性が出てくるので、だから、そうしてまさに。

江間：裏の裏の裏を読んでるから、そういう（笑)。

大澤：そうそう。だから、その関係に非常に私は興味があります。だから、人狼ゲームとか。すいません。それ、私の話は。

斉藤：いえいえ、面白いね。

大澤：そこに可能性を感じています。知能の課題として。

江間：何を、知能と定義するかみたいなところにもよってくる。何か結構、人工知能の研究者の方とお話しすると、そもそも何か、計算機でできるものは知能じゃないとかって、皆さん、おっしゃるじゃないですか。狭い意味では結構、もう人間とは何かみたいなのを本当に見ていきたいっていったときに、何かそんな単純なもんじゃないとか、それは知能ではないとか、内部のこと、よく分かってないとか、何かそういう感じで、新しいテーマを見つけてくフロンティアスピリッツにあふれた分野だっていうふうに。

斉藤：計算機科学のけん引役みたいなところはあったような気がする。つまり、よりパワーのある処理がしたいとかっていう、そういうのはあったと思うけど、でも、何か計算機でできたら知

能じゃないっていうのは何か結果論だし、まあまあ、だから、もう人工知能の対象から外してく。できたものから（笑）。

江間：できたらAIじゃない。

斉藤：数値積分はやれるんだから、もういいよとかっていう。そうやって、だんだん。それだけのことじゃないかしら。別にあんまり。

久木田：こないだ、まつもとゆきひろさんの講演を聞いたときに言ってたんですけど、最初はコンパイラのことをみんな、人工知能って言ってた。

大澤：ああ。

斉藤：コンパイラが人工知能？　そりゃあ、聞いたことないな、私。

大澤：そうですか。

斉藤：コンパイラが人工知能。

大澤：でも、本当に最初の頃は探索とか、あれがそうですね。

斉藤：サーチはそうだね。でも、計算機科学だよな。ツリーサーチとか。

大澤：今でいう、データベースの研究みたいなの、結構多い。当時のからスタートしてる。

斉藤：確かに。そう言ってたかもしれない。

大澤：皆さん、もう相違はないですけどね（笑）。だから、よく言われる、成功すると人工知能の名前が入るというやつだと思いますが。

服部：プロダクションシステムとかやってると、データベースの話も、それはAIの話の中で言いたくなるし、サーチもそうだよね、なんだけど、あれ自体もああいう結果に終わったので（笑）。個人的には、それ、サーチ、人工知能もやっぱりサーチだよねとは思うんだけれども、ただ、AI技術っていって、サーチとかっていわれると、さすがにちょっとそれは小さく言い過ぎだよねとかっていうふうに映る。

大澤：サーチは全てっていうふうな。でも、サーチって結構大きい課題って言われると、そうかなって気もしますが、今、でも、本質的に意味のあるサーチみたいな話だと、それこそ見た中から、停止の信号、持ってくるってのはそんなに言うほど簡単ではないっていうような。

久木田：囲碁っていうのはサーチの問題って言っていいんですか。囲碁、囲碁はサーチの問題？

大澤：ああ。でも、基本、回答は決まってるわけですから、理論上は。

斉藤：いや、ゲームはみんなっていうか、だから、ああいう類のゲームは、みんな、ツリーができるわけで、ゲーム木、そん中のサーチだよね。広い意味ではね。だけど、何かそんな全部は全然、サーチしてないし、ばかサーチしたら、全然、時間が足りないし、どんな計算機でも。だから、違う仕組みを使って人間もやってるし、計算機も違う仕組みでやるようになるし。サーチをね。オートマトンスタディーズに出てくる、アシュビーの、何だっけ。インテリジェンスアンプリファイアの話は、知能の本質はセレクションだって、確か言った。つまり選ぶこと。また選択肢があって、その中のどれがいいって選ぶっていうのが、抽象化して、単純化して、話をすると「ああ、知能なんだ」

みたいな。「そう？」みたいな（笑）。いや、でも、単純にものを買うときとか、カタログとかあれしたり、何か、最近だとオンラインでいろんな人の評判見て、じゃあ、どれにしようか。テレビ、大画面テレビ買うんだとかって、どれにしようかって。選ぶ動作だよね。そのときに一応、インテリジェンスが働いてるよね。嫁さん選ぶのも、もっと高度なインテリジェンスかもしれない。人間選ぶのは、もっと大変だよね。いや、人を採用するとかそういうのも全部、そうだけど、人を評価して、この人、この仕事できるかとかっていうのはかなり高度なインテリジェンスだよね。

大澤：あれですね。あと、インタラクションな系は、結局、価値観も途中で変わるし、どっち選んでも正解っていうこともあるじゃないですか。

斉藤：どっち選んでも正解。

大澤：つまり、選んだことによって価値観も変わるし、何かこう、そういうことってありませんか。

斉藤：選んだことで価値観が変わる？

大澤：そうそうそう。

斉藤：だから、変化してくってのはあるけど、分かんない。嫁さんの話の場合は、すごいゆっくりした変化じゃない？そんな急激には変わらないよね。だけど……。

服部：それ、社会性の話なんじゃない？

大澤：そうかなとか言って。でも、そうかもしれません。そうね。

江間：そろそろ、何か。

服部：そうですね。ちょっと多分、こういう話になってくと終わらなくなってくる。では、いったんここで。はい、すいません。大体、いつも話がだんだん乱れていく（笑）。ご質問、何か項目とかつくっても、大体、途中から外れて。はい。じゃあ、ちょっといったんここで区切らせていただいて、続きはもうちょっと滑らかなところで。どうもありがとうございました。

斉藤：いえ、どうもありがとうございました。

一同：ありがとうございました。

斉藤：私も楽しかったです。

（了）

AIUEO 1984.7.21

"Is it possible to represent symbolizer symbolically?"

斉藤康己（武蔵野通研）

状況設定

1. History of AIUEO （資料1 参照）
   AIは何をしてきたか？ 何が得られたか？
   program without theory は何の役にも立たない！

2. 8th-IJCAI at Karlsruhe, PANEL of Cognitive Modelling.
   "Under What Conditions Can a Machine Attribute Meaning to Symbols" (1983)
   Aaron Sloman, Drew McDermott, W.A. Woods, Brian Smith and Pat Hayes. （資料2 参照）

3. D. Marr "Artificial Intelligence - A Personal View"
   AI vol. 9, pp. 37-48. (1977)
   （資料なし）

   Type 1 Theory
   the isolation of a particular information processing problem,
   ☆ the formulation of a computational theory for it,
   the construction of an algorithm that implements it.
   a practical demonstration that the algorithm is successful.

   Type 2 Theory
   a problem (biological information processing problem) is solved by the simultaneous action of a considerable number of processes, whose interaction is its own simplest description
   ex. protein folding

4. J.A. Fodor の "The Modularity of Mind" （資料3）
   Fodor の input system が私の言う symbolizer に相当する。

5. 学習リンケージでの議論 （資料4）

   第1種と第2種の学習／知識
   cognitive science knowledge engineering

議論の骨

1. Symbolの魔力
Symbolで表現するとはどういうことか?

- 数の表現，代数式の表現　ex. 43, ab, 7 1/3
- 方程式による表現（物理法則）　ex. $F = G\frac{Mm}{r^2}$
- 化学反応の表現　ex. $HCl + NaOH \rightarrow NaCl + H_2O$
- バクテリオファージの形成　（資料 5）
  ↑
  これをsymbolicに計算機でどう表わすか?
- 脳の活動は?

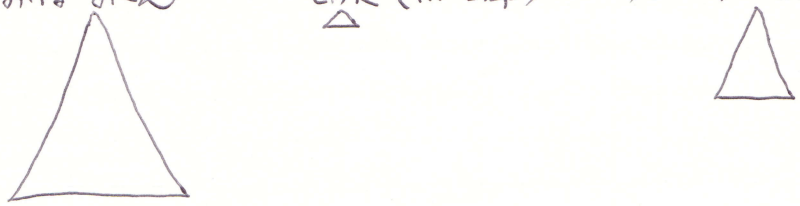
2. 脳は有限オートマトンだから必ず計算機で記述できる。
人間/脳は"変化"する有限オートマトンではないか?

原理的可能性と現実的実現性の間のギャップ.

3. 直接feedbackと間接feedback.　（資料 6）
|
interactionが重要に!

4. Symbolizerを計算機は持たない!

おばあさん　　CAR (in Lisp)　　犬の持つ主人のimage



5. よりむずかしいプロセス（進化 / 学習）
進化を方向づける"すりへる"ような必然性をどうやって
symbolの世界に持ち込むか?

6. symbolizerを記述しようとすると何がむずかしいか?
Non-modularity ?
外界とのinteractionを介して形成される → why?
何がタネか
何がinteractionで得られているか?

すり合わせ

資料2　8th. IJCAI proceedings pp.44-48　1984.7.21 AIUEO 担当 斉藤

PANEL DISCUSSION
UNDER WHAT CONDITIONS CAN A MACHINE ATTRIBUTE MEANINGS TO SYMBOLS

Aaron Sloman
(Panel Chairman)
University of Sussex

It is easy to produce arguments purporting to show that computers can never be conscious, have purposes, understand anything. One standard response is to claim that since computers can manipulate symbols they have the potential to support any kind of mental state or process, if appropriately programmed. The task is then to investigate what sort of programming will transform a symbol user into something with mental states. Much work in AI is concerned with specific subgoals of this investigation, taking it for granted that machines can manipulate symbols with meanings.

But can this be taken for granted? Certainly computers construct and transform patterns. We can interpret them as meaningful, but it does not follow that the computer does, any more than a filing cabinet interprets documents. If interpreting symbols already presuppose being conscious, having beliefs and goals, etc., then the ability to assign meanings to symbols cannot explain these other abilities.

One popular reply is to postulate sensors and motors. No mind or computational process can refer to particular events or objects in the physical world without being causally embedded in the world. Perhaps even reference to properties and relations (universals) requires such causal embedding. So if we find suitable causal relationships between internal processes and external events, then this would seem adequate to justify the claim that the internal processes use symbols with a meaning.

But is this necessary for every form of mentality, every meaningful use of symbols? Can't we imagine someone totally paralysed, blind, deaf, anaesthetised, yet conscious, thinking, even experiencing hallucinatory physical sensations - or at least solving problems about number theory? Or do we only think we can imagine this? People often assemble meaningful words into deceptively incoherent phrases, like "the view of the universe from outside time".

I claim that we should accept the anti-behaviourist intuition that the meaningful use of symbols can be purely internal. This is not just a semantic quibble about the meaning of "meaning". Rather, the point is that different sorts of internal computational architecture have profound implications for different sorts of mentality, independently of external relations. It is significant that computers have been designed in such a way as to support talk of machine-languages, instructions, addresses, conditionals, logical operations, arithmetic operations, and the like. The use of such terms in computer science does not depend on external relations. It is justified by the structure of internal processing. For instance, a computer can use numbers to count internal operations, or to count internal symbols, just as we use them externally.

The issues are surprisingly tricky. Given an uninterpreted computer you may be able to work out which symbols the machine interprets as booleans. But could you tell which one was taken as "true" and which as "false"? Can the distinction can only be made only in terms of a computational architecture rich enough to support a division between beliefs and goals? We don't notice the need for this normally because machines come with manuals which beg the question by telling us which operation is "and", which "or", etc.

We should explore the ability of different sorts of computational architectures to support such notions as reference, instruction, goal, description, truth, plan, desire etc. In particular, we need to understand the conditions under which a machine could have beliefs and goals relating only to internal states and processes, such as the belief that two lists have the same contents, or a goal like counting the number of even integers in an internal list. (The differences between beliefs and goals lie only in their roles in processing.) The relevant architecture need not be directly implemented in a physical machine. It could be a virtual machine implemented in a lower level machine with a simpler architecture. All this contradicts a philosophy which stresses the social, communicative, role of meaning.

So ordinary digital computers, suitably programmed, seem capable of supporting quite complex notions of meaning and mentality independently of external connections. Paradoxically, it is far less obvious how brains can support the self-referential abilities which enable ordinary computers to be described in mentalistic language. Do brains intrinsically require external connections for their use of symbols with meanings?

I have deliberately invited a biased panel to discuss these issues. When one side of a debate is so easy to argue for, we may hope to profit most by listening to the other side.

PANEL DISCUSSION
UNDER WHAT CONDITIONS CAN A MACHINE ATTRIBUTE MEANINGS TO SYMBOLS?

Drew McDermott
Yale University

1. In what sense do humans assign meanings to symbols?

2. In what sense do humans think they assign meanings to symbols?

3. Could machines assign meanings the way humans do?

4. Could machines think of their meaning-assignments the way humans do?

Many humans think that the first step in perception is sensation, and that larger percepts are made of atomic sensations. We now know that it is not necessary to appeal to sensation to explain perceptions. (You could arbitrarily label the first events in perception "sensations", but there is no compelling reason to.) It looks now as if we could design a complete vision system that had no intuitions whatever about "atomic feelings" occurring inside it. Why then do we think sensations play a role? How could we augment the vision system so it could think this about itself?

Similarly, we should expect in studying meaning (and other mental phenomena) to find discrepancies between what actually happens and the way we perceive it. One such discrepancy accompanies our overwhelming intuition that when we know that a symbol refers to something, we are "connecting" an abstract symbol to a concrete reality. I don't just have a formal theory about an individual named Reagan; I know who this name refers to.

But when we build robots, there is nothing inside the robot to actually "connect" a symbol to. In fact, the major intellectual achievement of computer science in western culture is to demonstrate that a device can manipulate symbols "correctly" without knowing what they mean. Previlusly, mind-theorists who tried to build theories based on symbol manipulation kept stumbling over the homunculus, required to understand the symbols being manipulated.

Does this mean that a robot does not succeed in referring to the world at all? No, as the following sketch should demonstrate.

A robot can be modelled as one or more formal systems, connected to the world by way of sensors and effectors. They will be formal, systems in that they operate by applying elementary operations to uninterpreted symbol structures, deriving new uninterpreted symbol structures. The sensors create the first symbol structures; the effectors receive some of the inferred symbol structures, and execute them as instructions.

This works because the formal system has an interpretation, which matches the real world closely. I.e. there is an assignment of real-world entities, properties, and propositions to the symbol structures of the formal system, with the following properties:

* Input soundness: the sensor apparatus is constructed so that the symbolic structures it constructs usually correspond (under the interpretation) to actual states of affairs; that is, they are usually true.

* Inferential soundness: the formal system is constructed so that it tends to infer true symbolic structures from other true symbolic structures.

* Output soundness: the effector apparatus is constructed so that it tends to perform the actions that the interpretation assigns to the symbol structures given to it.

For example, when a car is coming, the robot's sensors put an expression (CAR-COMING) into the formal system, which infers (SHOULD-DO (JUMP)), and sends (JUMP) to the effectors, which actually causes a jump.

Having an interpretation, does not mean that the formal system has access to something, but that there exists an interpretation in the mathematical sense. It is futile to ask for more. If the robot had access to a thing purporting to "contain the meanings" of its symbols, then either this thing would be just another formal system, and wouldn't contain the meanings after all; or the robot would not be a Turing machine, but some more mystical entity. If we want to stick with the formal robots that have revolutionized our thinking about the mind, we must place the interpretations of their symbols outside their heads.

So the interpretation can play no functional role in how the robot works. It is simply an analytical tool, to explain how it works.

What if there is more than one interpretation of the formal system? After all, any interesting system will have an infinite number of different

models. But most of them will fail to satisfy the three soundness conditions. There is an interpretation in which (CAR-COMING) means "Rice pudding present", and (SHOULD-DO (JUMP)) means "My mother is standing on her head", but the sensors, effectors, and inference machinery do not work correctly under this interpretation. I expect that for any robot, there is a "standard sound interpretation" that does satisfy the soundness conditions in our world. Since these conditions are stated as engineering precepts ("component X usually operates according to the interpretation"), there are probably lots of unimportant variants of the sound interpretation, but I will ignore this issue.

I have now answered my original questions 1 and 3, how meanings actually work. We must now ask how humans and robots might think about how they work. First, let me acknowledge that there is a mystery here about why humans have any opinion at all about whether they assign meanings to symbols. The answer might be that humans are just naturally inquisitive, and have opinions about everything, but perhaps there is some special reason to have opinions of certain kinds about oneself.

The symbols people think they manipulate are not those in the robot's formal system. When people think of symbols, they think of words, names, mathematics, and road signs. Suppose a human, to be specific, Edwin Meese, is mediating on the meaning of "Reagan". He knows that this symbol refers to his boss, person he sees often. He believes he can think about Reagan any time he wants, and that this name refers to that person.

The truth is that for Meese to think about Reagan is for him to manipulate symbol structures that refer to Reagan in his standard sound interpretation. Ironically, for him to think about the name "Reagan" is to manipulate symbol structures in much the same way. (Since the name is a social object about which various things are known, e.g., it's pronounced differently from the name of the Treasury Secretary.) The idea that the one object refers to the other is a third symbol structure, used mainly by natural-language modules.

The mind conceals such facts about itself. (It is usually pretty easy to see why, but that isn't my main topic.) It tells itself that to think about Reagan is to "directly apprehend" him. You can think about an object if you are "acquainted" with it, if you know "which object it is".

This is contrasted with another situation we have all been in, where we know someone (or something) by name only. Suppose someone has been marooned on a desert island for twenty years, and, having returned, hears people blame someone named "Reagan" for all our troubles. Obviously, this is someone he ought to know about; he ought to know who the name refers to. Somehow, just knowing that Reagan is "the person everyone is blaming for our troubles" is inadequate. After finding several more facts, and seeing Reagan on television, he begins to feel tht he is "knows who Reagan is", that he can think about him any time he wants.

People feel a sharp difference between only knowing something's name and knowing the object directly. I think this is an illusion; in reality, one accumulates information about an object gradually. There is no well defined point at which one is "really" acquainted with it. The sharp feeling is akin to a sharp feeling of hunger; there is no qualitative difference between an empty stomach and a full one, but it feels like there is; if it didn't, you wouldn't work so hard at finding food. In the case of acquaintance, you need a reason to work hard to gather information.

Whatever the source of this feeling, it leads to disbelief that all knowing is mediated by a formal system. If all you have is symbols, then you aren't "really" acquainted with anything, and you don't "really" understand anything. In fact, the sketch I started with explains quite satisfactorily how a purely formal system can nonetheless deal with the real world, and, in a certain sense, have its symbols mean things in that world. Being "directly connected", or "knowing what the symbols mean", plays no role.

This answers my second question, How do people think they assign meanings to symbols? The fourth question, How might we get machines to think about themselves this way?, I will leave unanswered.

IJCAI SYMPOSIUM
UNDER WHAT CONDITIONS CAN A MACHINE USE SYMBOLS WITH MEANING?

W. A. Woods
Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, MA 02238

Most existing computer systems use representations that have meaning to their designers, but do not stand in any relationship to their intended meaning whereby one could say that they had that meaning to the machine itself. I will be concerned here with the question of when symbols* actually have meaning to the machine. This question presupposes the answer to a prior question: why should one maintain a distinction between the meaning of a symbol and rules that simply imply its applicability or inapplicability? I will attempt first to answer this question, from which an answer to the former will follow. Briefly, I will argue that the distinction is necessary in order for a system to make proper adjustments to its knowledge base as a result of experience.

In a previous work [1], I have argued that the meaning of a symbol resides in an abstract procedure, not necessarily executable, linking the symbolic expression to the physical world through the (computational / inferential) operations of a physical interpreter operating on a combination of internal representations and sensory/motor connections to the world. Moreover, I argued that this meaning specification (the <u>meaning function</u>) is distinct from the procedures that one ordinarily uses to conclude the truth of a description in everyday situations (<u>the recognition functions</u>).

I will assume here that both meaning and recognition functions can be represented by collections of rules of the form "p implies q", provided that a distinction is maintained by the system between those rules that it applies as part of the meaning (i.e., "meaning postulates") and other implications between symbols (which I will refer to as <u>hypotheses</u>).

If a system is to try to develop a model of the world in which it operates, then it needs to record rules of the form "p implies q" where, in the simplest case, p denotes some observable attribute of the world and q denotes some other observable attribute. The utility of such a rule lies in the ability to conclude q without actually perceiving it. Let us call p a <u>sensory</u> symbol (for a system) if there are processes in the world that (physically) cause the system to conclude p or to conclude not p. I will take the physical processes which operate in this way as determining the meanings of such sensory symbols.

A system with no sensory symbols could not be expected to acquire a model of the world except through the intervention of an external agent (i.e., a programmer). In this case the meanings of its symbols would reside solely in the head of the programmer, and the system would be open to different interpretations by different external observers, however consistent its internal operations were with some particular interpretation.

Consider a system that has both sensory and nonsensory symbols and that hypothesizes rules of the form "p implies q" under some circumstances. A validated hypothesis "p implies q" is useful to the extent that one knows under what conditions to invoke the hypothesis (i.e., one knows the meaning of p) and one knows what claims are being made thereby (i.e., one knows the meaning of q). Moreover, in order to determine the validity of the hypothesis, one needs to be able to determine the truth and falsity of p and q in most circumstances.

Suppose that this system encounters a violation to a pair of rules "p implies q" and "q implies r" by perceiving that p is true and r is not. If q is a sensory symbol and its truth can be determined by perception in the situation in question, then the appropriate response is simple -- either the "p implies q" rule or the "q implies r" rule should be modified, depending on whether q is false or true. If q is not a sensory symbol, the system needs some criterion to decide which of the two rules to modify. Since one rule implies q is true, while the other implies it is false, what the symbol q means (either to the system or to an external agent) becomes important at this point in deciding what to give up.

To claim that the symbols in a machine have meaning is to claim that above and beyond the rules and procedures that lead them to be deduced by the system, there is some further claim about the state of the world that is being made by virtue of such a deduction. That is, regardless of one's reasons for believing p, there is some claim about the world that p makes, which may be either true or

---

*I will use the word "symbol" here as a generic for the entire class of symbolic expressions used by the system, including designating expressions, predicates, propositions, etc.

false. Without such a distinction, a system would be unable to formulate a predictive model of its world, and could at best keep a record of its experiences; no deduction that it made could ever claim anything more than that it had made the deduction.

One of the main benefits from having a notion of the claim made about the world by a symbol (i.e., its meaning) is that if one encounters a sequence of deductions from some sensory symbol p to some sensory symbol q, while perceiving that p is true and q is false, then one can attempt to localize the erroneous rule by examining the truth of the intermediate symbols involved in the deduction. In particular if some intermediate symbol s is determined to be true (by virtue of its meaning) then the deduction leading from p to s may be left intact and attention directed to finding the error in the deduction from s to q (and conversely if s is determined to be false).

Without some such criterion for localizing the error, any one of the intermediate rules could be changed to eliminate the inconsistency in the momentary situation, but with little confidence of the effect on other hypotheses (perhaps already strongly verified). That is, it would be inappropriate to change the meaning rules (and thereby the meaning) of any of the symbols already involved in partially verified hypotheses, thereby changing the claim made by those hypotheses and either invalidating them or weakening their previous validation.

One can imagine situations in which the minimal change to an entire belief system in order to bring it into line with some body of observations is to make a change in the meaning of some symbol, thereby eliminating inaccuracies in, or increasing the coverage of, existing hypotheses. However, such a change would have to come from a complex analysis of a large portion of the system's knowledge base. This would be a revolutionary rather than evolutionary modification and would not be an appropriate mechanism for performing routine, incremental modifications to knowledge.

If one considers such modifications in an internal language of thought, coupled to symbols in the external vocabulary by connections that can themselves be learned and modified, then any such revolutionary change could also be effected with no changes in the meanings of internal symbols by merely moving the affected hypotheses (and appropriate external symbols) to new (internal) symbols with the desired meanings rather than changing the meanings of the old ones. Thus, in the internal language there appears to be no need to ever change the meaning of a symbol.

An obvious question, at this point, is whether an analysis along the above lines can apply to natural kind terms as well as to terms with obvious necessary and sufficient conditions. I believe that this is the case, although the story now becomes a little more complex. In this case, the question of interest about the meaning of a symbol has to do with what natural kind the system has got hold of when it forms a hypothesis. (In general a satisfactory account of natural kinds must allow the meaning rules to leave "gaps" in which the system has no intentions as to whether some hypothetical situation embodies an instance of a symbol or not, and must permit meanings to be extended in certain circumstances by narrowing the gap -- see [1] for further discussion.)

Consider the hypothesis "snakes are harmless" expressed in New England, where there are virtually no poisonous snakes. One may ask whether the meaning of the internal symbol for snakes in this case refers to only (nonpoisonous) snakes of the kind that live in New England, or to all kinds of snakes. (I assume that both these classes are natural kinds). The answer, I would argue, depends on the intentions of the person when the concept snake was formulated and the preservation of those intentions in the meaning rules for the symbol (i.e., whether the scope of the meaning rules was sufficiently broad to cover all snakes or included enough specific characterization of the range of snakes experienced to rule out the poisonous kinds that exist elsewhere).

Once again, the meaning intended and encoded in the meaning rules will determine the reaction to encountering an anomalous situation -- in this case, whether to attach the external word "snake" to a new internal concept which includes poisonous snakes, or to subclassify the the internal concept and move the "snakes are harmless" hypothesis to a more specific subclass.

If we take the above simplified story as the reason for maintaining a distinction between the meaning of a symbol and a mere collection of conditions that imply it or that it implies, then it follows that for symbols to have meaning to a system, there must be sensory symbols to which the nonsensory symbols are related by a distinguished set of meaning rules or some equivalent mechanism. These rules together with their interpreter and the physical processes that govern the senses determine the meanings of the system's symbols.

By this account, a system that read stories and answered questions about them, based solely on the manipulation of internal representations with no experiential base, would, from its own perspective, be manipulating meaningless symbols. Similarly a pocket calculator would be a manipulator of symbols whose meanings were externally attributed (even though its internal structure does faithfully model the intended interpretation). On the other hand, a simple robot creature with primitive perceptions can be said to have meanings for its symbols that are not externally imposed if its symbols are tied to those perceptions by some distinguished mechanism equivalent to meaning rules.

Reference

[1] Woods, W.A., "Procedural Semantics as a Theory of Meaning," in A. Joshi, I. Sag, and B. Webber (eds.), Elements of Discourse Understanding, Cambridge University Press, 1981, pp 300-334.

情一室　談話会　資料3 (1/2)　1984年5月11日 1h30m-

『The Modularity of Mind』 by Jerry A. Fodor

～ An Essay on Faculty Psychology ～

résumé par Yasuki SAITOH

A Bradford Book, The MIT Press, 145 pages (1983).

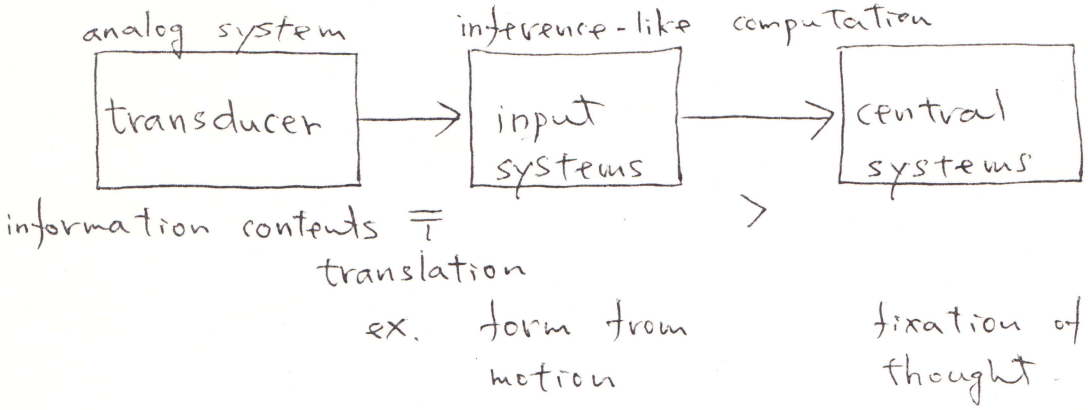
- modularity ← modular ← module =
  a part of something that can function independently of the other parts.
- faculty = the power or ability to do something particular.

全体は以下の5部構成になっている。

Part I Four Accounts of Mental Structure
　II A Functional Taxonomy of Cognitive Mechanisms
　III Input Systems as Modules
　IV Central Systems
　V Caveats and Conclusions

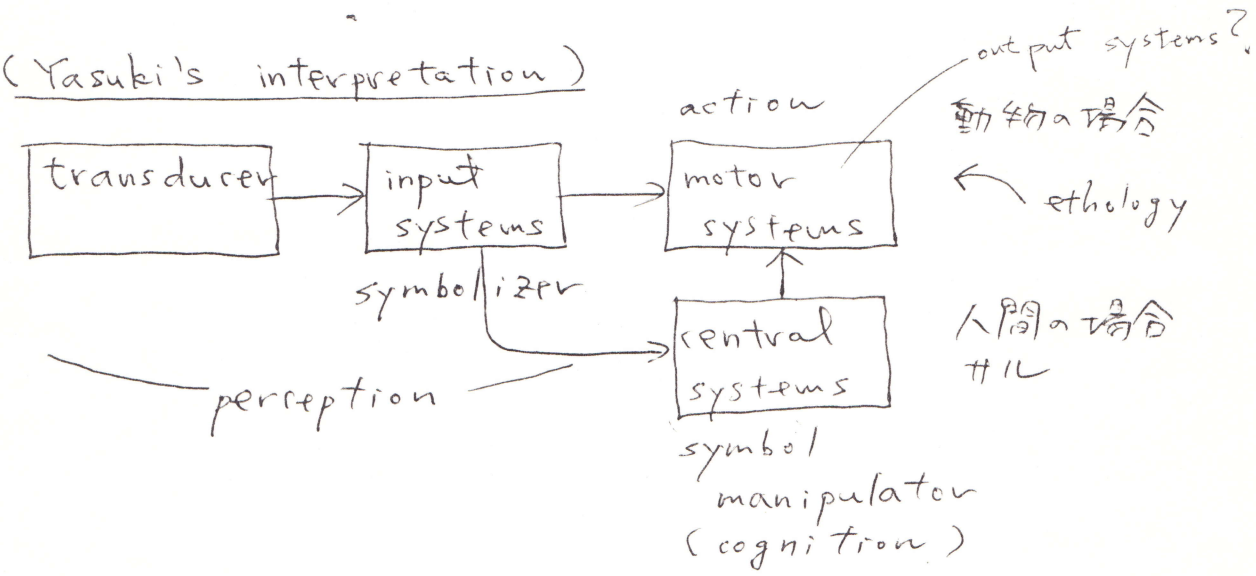
第1部で、心の構造に関するいくつかの説を紹介したあとで、その中の1つ Gall の vertical faculties の説に若干変更を加える形で Cognitive System の Taxonomy を与える(第2部)。続いて、第3部では、その中の input system に焦点をあてて、その性質を列挙し、説明する。input system は domain specific, modular, informationally encapsulated, などなどの性質を持つ。これと全く反対の性質を持つものとして第4部で、Central System が取りあげられる。Central System を特徴づける形容詞は、Quineian, isotropic である。最後に第5部で、epistemically bounded / unbounded の議論と modularity thesis の関係に関して警告を述べ、Cognitive Science が研究対象とできるのは input systems のみという結論で終る。

PART II A Functional Taxonomy of Cognitive Mechanisms



- central system が computation を行えるような形に world を represent するのが input systems の役割.

- (Yasuki's interpretation)



- 計算機の場合は、人間が input systems の役割をはたしている。
  programmer

資料 4 (1/2)

学習リンケージ用資料　　1984.4.19

What is learning?
Man's learning and Machine's learning

情一　斉藤康己

Key word - symbolizer

☆My Definition:

外界とのインタラクションを通じて，有機体(必ずしも生物である必要はない)がその構造を外界にマッチするように変化させ，よりうまく行動するようになることが学習である。

☆Classification of Learning:

学習は，それに要する時間，及び，その内容の観点から大きく3つに分類できると思う。

1：大変時間のかかる学習 ⟶ 生物の進化，社会の発展

2：ある程度時間のかかる学習 ⟶ 言語の習得，常識の形成
Symbolizer の確立
↖ J. Fodor の言う "input systems" の形成
cf. "The Modularity of Mi

3：あまり時間のかからない学習 ⟶ 言語による知識の獲得
一夜づけ
電話番号のごろによる記憶

☆Machine's learning:

現在の計算機は Symbolizer を持たないので，学習には大変な困難が伴なう。
↙ IJCAI-83
"Under What Conditions Can a Machine Attribute Meaning to Symbols?"
もしソフトウェアやそれを開発する人，使う人まで含めて，1つのシステム（有機体）とみなせば，1のレベルの学習は存在すると思われる。 ⟶ 人間を介してのソフトウェアの進化.

☆Conclusion:

以上の考察からAIの研究として学習を取り上げ，計算機とのからみで何か研究しようと思うと，Symbolizer の部分がネックになるし，又それが一番 attack すべき問題であると考える。

資料4(2/2)

84.7.4

学習リンケージ

これまでのまとめ 1

視覚G 今田

1. 学習とは?

未定義のまま出発

2. 学習の種類

① 第0種の学習 (第0種の知識)
才1種, 才2種の学習を行なうための手段を提供する

② 第1種の学習 (第1種の知識)
画像・音声等, 自然言語化できないものに関する知識の習得

③ 第2種の学習 (第2種の知識) basic category
自然言語等に基づく論理的な知識の習得

3. 学習時の問題点

① 汎化と分化
② 類推 (ルールの汎化・分化)
③ 概念化
④ 帰納
⑤ 知識の使用範囲の限定

4. 学習の目的

学習によって得た(記憶した)知識の中で, ある外部状況に最も適した知識を参照できる or 出力できること。

要するに, 学習は, その知識を使うときのことを考えて行なうべし。

5. 知識を使うときの状況

幾つかの条件 ⟶ 知識の範囲の限定 ⟶ 最適知識の選定
(外部状況)

Production rule の前処理版

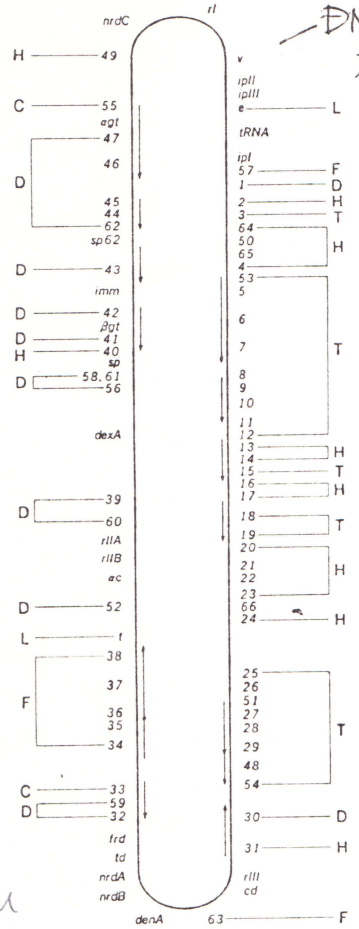
資料5 (1/2)

飯野てつふ
分子構造の遺伝学
の7章

4

1.3 バクテリオ・ファージ T4の形成

DNA 分子量 約1億
ヌクレオチド対 約17万
遺伝子総数は
150コぐらい



長さ50μ

図 7-5 T4ファージ染色体上における各種遺伝子の分布。
斜体の文字は遺伝子記号を表わす。Hは頭部，Tは尾部，Fは尾毛，DはDNA合成，Cは調節，Lは溶菌をそれぞれ支配する遺伝子である。また矢印はオペロン単位と転写の方向を示す。

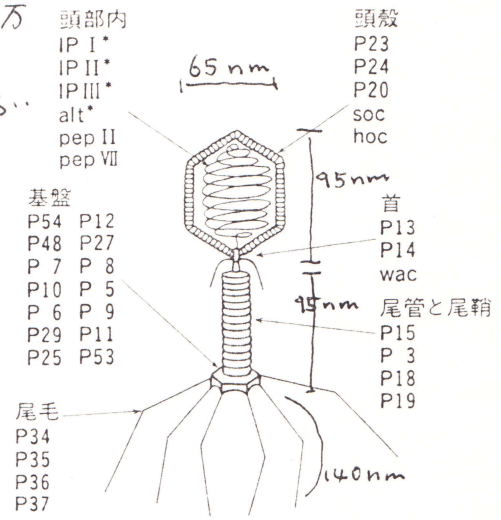


図 7-1 バクテリオファージ T4 の構造タンパク質。
頭部全体で14(首も含まれる)，尾部に18，尾毛に4，計36種類(少なくとも)が存在する。DNAは頭部内に存在する(柳田充弘，化学と生物 15: 140(1977)より)。

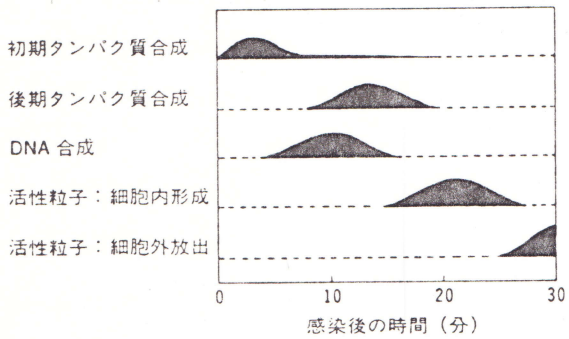
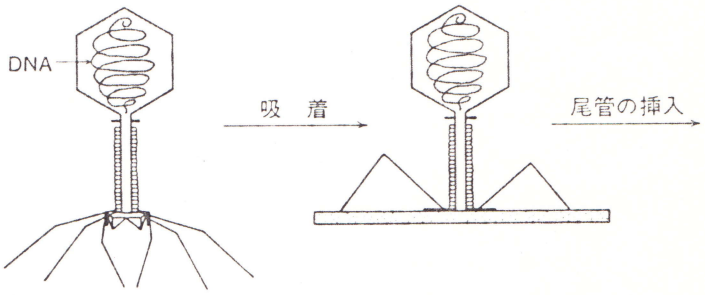


図 7-4 T4ファージの増殖過程における反応の継時的変化



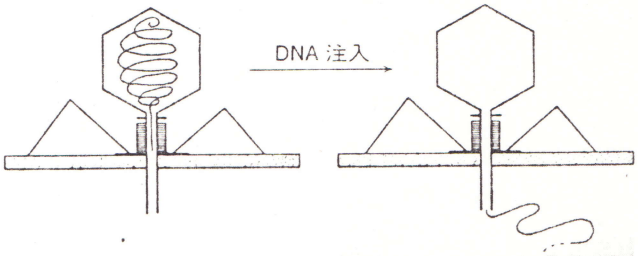


図 7-3 T4ファージが大腸菌に感染する際の粒子の構造の変化。

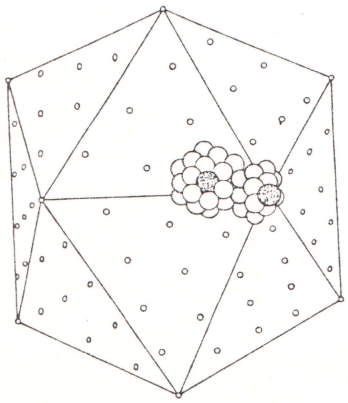


図 7-6 T4ファージ頭殻の成分タンパク質の配置。
面上の集合単位1個と角の集合単位1個についてモノマーの配置を示してある。○はP23*，◉はhoc，○はsoc，∘はその他の集合単位のhocの位置を示す。

○部分集合体の段階的集合

並列構築
(parallel assembly)
↑
突然変異体の
研究から

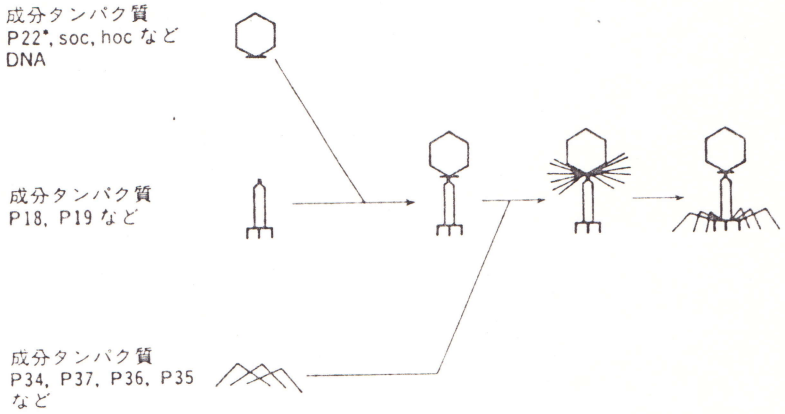


図 7-11　T4 ファージの部分構造体の集合による粒子形成の過程。

○頭部形成（遺伝だけで形態がきまるというのは全くの誤りだ！）

scaffolding structure

殻タンパク質の変化

分子量の変化	
P23	56,000
P23*	46,500
P24	45,000
P24*	43,500

切断用の酵素タンパク質は P21* である。
← P21

分子加工
(molecular processing)

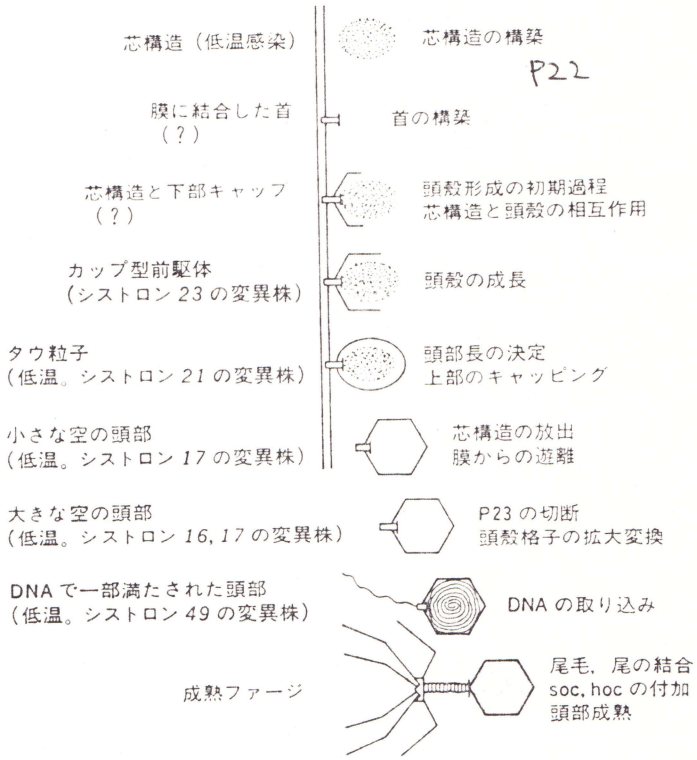


図 7-12　T4 ファージの頭部形成における構造変化。
かっこ内は宿主細胞内に蓄積される条件を示す（柳田充弘・宝谷紘一，分子集合体の構造変換と構造形成，"生体における分子識別"(1979)，共立出版．より）。

DNAの挿入と頭殻の成熟
pA（タンパク質），スペルミジン，ATP，Mg^{2+} → DNAが中へ入る。
DNA切断酵素

○尾管と尾鞘の構築
順序性，自己集合反応　P19モノマー 144個 ← 分子歪みの集積

○尾毛の形成　P34×2
P37×2 + P36×2 + P35 →

資料 6 (Y1)

1984.7.11
(水)

NL, PL, Logic and AI
(水曜会用アジビラ)

情一 斉藤康己

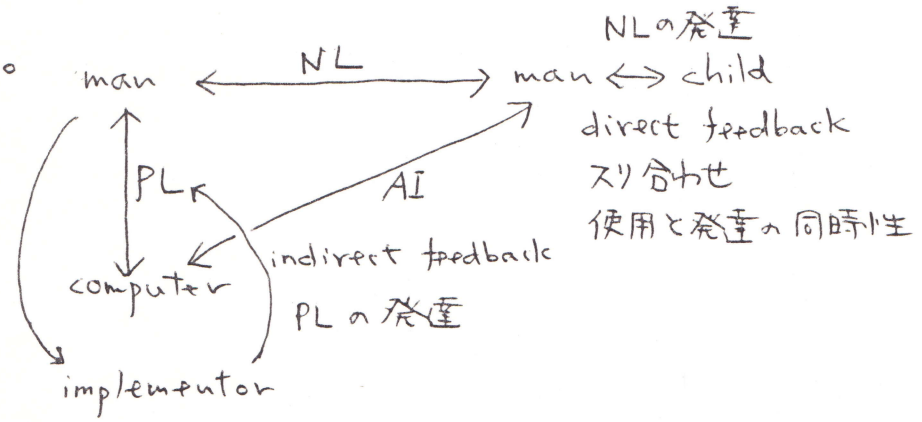
motivation

- AIの基本的問題の1つとして dog recognizer を記述することがある。
  grand mother でも triangle でもよい
  (Fodor によれば input system, Yasuki によれば symbolizer)

- いまだかつて誰も dog recognizer を記述できていない。
  その原因は対象の本質的複雑さにある。
  あるいは記述用の道具(言語)の稚拙さにある。

- しかし、脳は有限オートマトンだから原理的には必ず記述できるはずである。 → 道具(言語)のどこが不備なのか?
  ↓ として考えられるのは
  NL Natural Language
  PL Programming Language
  Logic
  これらの関係及び限界は?

- 道具のみを取り出して、それのどこが不備なのかを考えるのではなく、道具とそれを使う Agent のペア全体として何が足りないのかと考えるべきである。

- 

- PLでは1段上の level までしか組み上げられない。(人間の能力の限界)
  NLでも同じことか? (set of vocabulary がないと考えられない)

- Logic の分野で得られた限界(レーヴェンハイム/スコーレム/クルスキーの定理)がそのまま NL にも当てはまるか?
- "Under What Conditions Can a Machine Attribute Meaning to Symbols?"

1984.7.21 AIUEO 担当 斉藤 資料1

type aiueo.prehistory.l
1976-1977 at Essex

[Language]
Charniak: "Jack and Janet in Search of a theory of knowledge"
Charniak: "Organization and Inference in a frame-like system of common sense knowledge"
Schank: "Identification of Conceptualizations Underlying Natural Language"
Schank: "Using knowledge to understand (SAM)"
Wilks: "Summary of Natural Language tutorial lecture"
Rieger: "The commonsense algorithms as a basis for computer models of ..."
[Vision]
Gregory: "Seeing" and "Art & Reality" from Eye and Brain
Minsky,Papert: "Progress Report on AI"
Winston: "The MIT Robot"
Marr: "On the purpose of low-level vision" (seeing under things!)
Weir: "The perception of motion: Actions, Motives & Feelings"
Navatia,Binford: "Description and Recognition of curved objects"
Marr,Nishihara: "spacial Disposition of axes in a generalized cylinder representation"
[Planning]
Waldinger: "Achieving several goals simultaneously"
Sacerdoti: "The nonlinear nature of plans"
[Programming]
Davis,King: "An overview of Production Systems"
Mcdermott,Sussman: "Why conniving is better than planning?"
Green,Barstow: "Some rules for the automatic synthesis of programs"
Hardy: "Synthesis of LISP functions from examples"
Bornat,Waldinger: "Does AI programming really have to be like knitting with spaghetti?"
[Frame theory]
Minsky: "A framework for representing knowledge"
Winograd: "frame Representations and the Declarative/Procedural Controversy"
Bobrow et al.: "GUS, A frame-driven dialog system"
Bobrow,Winograd: "An overview of KRL, a knowledge representation language"
Bornat,Wielinga: "An underview of KRL, - is it a P.L.?"
[Semantic Net]
Woods: "What's in a link?"
Schubert: "Extending the expressive power of semantic networks"
[Geometry theorem provers etc.]

History of AIUEO compiled by Yasuki SAITOH

(Original members are Saito, Shirai, Nakashima, and Katagiri)
1977

12-Dec-77(Mon)10-12 Saito
M.Minsky, "Plain Talk about Neurodevelopmental Epistemology",
5th IJCAI, pp.1083- (1977).
19-Dec-77(Mon) ?

1978

23-Jan-78(Mon)10-12 Shirai
Roger C. Schank, "Identification of Conceptualizations Underlying Natural Language",
Chapter 5 of Computer Models of Thought and Language, (1973).
13-Feb-78(Mon)10-12 Nakashima
Sussman & McDermott, "Why conniving is better than planning?",
Carl Hewitt, "PLANNER: A language for proving theorems in Robots",
1st IJCAI, pp.295-301, (1969).
3-Apr-78(Mon)10-12 Katagiri
Bobrow, Winograd, "An overview of KRL, a knowledge representation language"
28-Apr-78(Fri)13-15 Saito
David Marr, "A Theory of cerebellar cortex",
Journal of Physiology vol.202, pp.437-470, (1969).
29-May-78(Mon)13-15 Nakashima? or Shirai?
?
12-Jun-78(Mon)13-15 Suzuki
J.L.Peterson, "Petri Nets",
Computing Surveys vol.9, No.3(1977)
26-Jun-78(Fri)13-15 Katagiri
D.Rumelhart & D.A.Norman, "The Active Structural Network",
chapter 2 of Explorations in Cognition (1975).
7-Jul-78(Fri)13-15 Shirai
T.Winograd, Lecture 1 of Winograd's five lectures on AI.
21-Jul-78(Fri)10-12 Saito
Douglas B. Lenat, "Automated Theory Formation in Mathematics",
5th IJCAI, pp.833-842, (1977).
Douglas B. Lenat, "The Ubiquity of Discovery",
AI vol.9, pp.257-285, (1978).
(5-Aug-78 Nakashima left Japan.)
6-Oct-78(Fri)13-15 Suzuki
C.Hewitt & H.Baker, "Laws for communicating parallel processes",
Information Processing '77 IFIP, pp.987-992, (1977).
C.Hewitt & H.Baker, "Actors and Continuous Functionals", Draft.
27-Oct-78(Fri)13-15 Yokoo?
Patrick H. Winston, "Learning by Creating Transfer Frames",
AI vol.10, pp.147-172, (1978).
17-Nov-78(Fri)13-15 Katagiri?
Kenneth Kahn & G.Anthony Gorry, "Mechanizing Temporal Knowledge",
AI vol.9, pp.87-108, (1977).
1-Dec-78(Fri)13-15 Saito
R.J.D.Power and H.C.Longuet-Higgins,F.R.S.,
"Learning to count: a computational model of language acquisition",
Proc. R.Soc.Lond.B. 200, pp.391-417 (1978).

1979

12-Jan-79(Fri) -> 19-Jan-(Fri) Saito
R. L. Gregory, "Will Seeing Machines have illusions?",

Machine Intelligence ?, pp.169-177, (?).
9-Mar-79(Fri)?
around here the following book was intensively read.
rotation is Shirai Suzuki Yokoo Katagiri and Saito
Margaret Boden, "Artificial Intelligence and Natural Man",
The Harvester Press, Hassocks (1977).
Part I Introduction
1. What Is Artificial Intelligence?
Part II The Personal Dimension
2. Artificial Neurosis
3. Function and Failure in the Neurotic Program
4. Personal Politics and Ideology Machines
Part III Language and Understanding
5. Responding to Language
6. Intelligence in Understanding
7. Sense and Semantics
Part IV The Visual World
8. Adding the Third Dimension
9. Glimpses of the Real World
Part V New Thoughts from Old
10. Learnig
11. Creativity
12. Problem Solving
Part VI The Relevance of Artificial Intelligence
13. Psychological Implications
14. Philosophical Issues
15. Social Significance

17-Mar-79(Sat)10-12
12-May-79(Sat)15-17 -> 19-May-(Sat)15-17
2-Jun-79(Sat)15-17 Shirai
??LUNAR chapters 1-6
"Semantics and Quantification in Natural Language Question Answering",
Advances in Computers vol.17, pp.1-87, (1978).
23-Jun-79(Sat)14-16 Yokoo
ditto
(29-Jun-79 Nakashima came back)
(colby AI&NM around here, parallel ?)
7-Jul-79(Sat)14-16 Saito
J. Backus, "Can Programming be liberated from the von Neumann style?
A functional style and its algebra of programs",
CACM vol.21, no.8, (1978).
21-Jul-79(Sat)14-16
(Nakashima OWL + LMS around here)
1-Sep-79(Sat)14-16 Nakashima
Jon Doyle, "A Glimpse of Truth Maintenance System",
6th IJCAI, pp.232-237, (1979).
13-Oct-79(Sat)15-17 Saito
?
27-Oct-79(Sat)15-17 Katagiri
R.C.Schank, M.Lebowitz, and L.Birnbaum,
"Parsing Directly into knowledge structures",
6th IJCAI, pp.772-777, (1979).
17-Nov-79(Sat)15-17 Shirai
Wendy Lehnert & Yorick Wilks, "A Critical Perspective on KRL",
Cognitive Science vol.3, pp.1-28, (1979).
8-Dec-79(Sat)14-16 Yokoo
Daniel G. Bobrow & Terry Winograd, "KRL: Another Perspective",
Cognitive Science vol.3, pp.29-42, (1979).
22-Dec-79(Sat)14-16 Suzuki
T.Winograd, "Beyond Programming Languages",
CACM vol.22, no.7, pp.391-401 (July 1979).

1980

19-Jan-80(Sat)14-16 Nakashima
Jerry R. Hobbs, "Coherence and Coreference",
Technical Note 168, SRI Artificial Intelligence Center (August 1978).
2-Feb-80(Sat) 14-16 Katagiri
P.R.Cohen, "Elements of a plan-based theory of speech acts",
Cognitive Science vol.3, pp.177-212 (1979).
23-Feb-80(Sat)14-16 Saito
Douglas R. Hofstadter,
"Godel, Escher, Bach: Eternal Golden Braid", Chapter XI Brains and Thoughts,
Basic Books Inc. Publisher, (1979).
8-Mar-80(Sat) 14-16 Akiyama
Yasunaga, "Bunretsubyo-shojyokikou-nikansuru-ichi-kasetsu".
5-Apr-80(Sat) 14-16 Katagiri?
D.McDermott, "Planning and Acting",
Cognitive Science vol.2, pp71-109(1978).
19-Apr-80(Sat)13-15 Saito
Yuichiro Anzai and Herbert A. Simon, "The Theory of Learning by Doing",
Psychological Review vol.86, no.2, pp.124-140, (1979).
10-May-80(Sat)14-16 Shirai
R.Jackendoff, "Grammar as Evidence for Conceptual Structure".
24-May-80(Sat)14-16 Suzuki (Enter Aida)
Johan de Kleer, "The Origin and Resolution of Ambiguities in Causal Arguments",
6th IJCAI, pp.197-203, (1979).
7-Jun-80(Sat)14-16 Nakashima
R.Waldinger, "Achieving several goals simultaneously",
Stanford AI Center Technical Note 107.
21-Jun-80(Sat)13-15 Yokoo?
Steve Rosenberg & Bruce Roberts, "Coreference in a frame database",
6th IJCAI, pp.729-737, (1979).
Steven Rosenberg, "Reasoning in incomplete domains",
6th IJCAI, pp.734-737, (1979).
5-Jul-80(Sat)14-16 Akiyama
Kawamura, "Kotenteki-jyokenzuke-no-shinkei-kikou",
Shinkei-shinpo vol.22, no.5,(1978).
26-Jul-80(Sat)14-16 Katagiri
Hubert L. Dreyfus, "A Framework for misrepresenting Knowledge",
chapter 6 of Philosophical Perspective in Artificial Intelligence.

9-Aug-80(Sat)..11-Aug-80(Mon) AI summer seminar
Aaron Sloman, "The Computer Revolution in Philosophy:
Philosophy, Science, and Models of Mind",
The Harvester Press (1978).
1. Introduction and Overview
PART One Methodological Preliminaries
2. What are the Aims of Science?
3. Science and Philosophy
4. What is conceptural analysis?
5. Are Computers really relevant?
PART Two Mechanisms
6. Sketch of an Intelligent Mechanism
7. Intuition and Analogical Reasoning
8. On Learning about Numbers: Some Problems and Speculations
9. Perception as a computational process
10. Conclusion: AI and philosophical problems

6-Sep-80(Sat)14-16 Katagiri
Terry Winograd, "Extended Inference Modes in Reasoning by Computer Systems",
AI vol.13, no.1,2, pp.5-26, (April 1980). Special Issue on Non-Monotonic Logic

20-Sep-80(Sat)15-17 Saito
Drew McDermott and Jon Doyle, "Non-Monotonic Logic I",
AI vol.13, no.1,2, pp.41-72, (April 1980).

11-Oct-80(Sat)14-16 Saito
ditto

1-Nov-80(Sat)14-16 Shirai
R. Reiter, "A Logic for Default Reasoning",
AI vol.13, no.1,2, pp.81-132, (April 1980).

22-Nov-80(Sat)14-16 Shirai, Suzuki
R. Reiter, "A Logic for Default Reasoning",
AI vol.13, no.1,2, pp.81-132, (April 1980).
John McCarthy, "Circumscription - A Form of Non-Monotonic Reasoning",
AI vol.13, no.1,2, pp.27-39, (April 1980).

13-Dec-80(Sat)14-16 Nakashima
Drew McDermott, "The PROLOG Phenomenon",
SIGART pp.16-20, (July 1980).

1981

24-Jan-81(Sat)14-16 Matsukata
Masami Hagiya, "Yoiko no Hyperlisp"
Masahiko Sato, "Theory of Symbolic Expressions",
Technical Report 80-16, (Sep 1980).

21-Feb-81(Sat)14-16 Mizoguchi
Patrick H. Winston, "Learning and Reasoning by Analogy",
CACM vol.23, no.12, pp.689-703, (December 1980).

7-Mar-81(Sat)14-16 Aida
Keith L. Clark, "Negation as Failure",
from Logic & Date Base

28-Mar-81(Sat)14-16 Suzuki
Harold Boley, "A Preliminary survey of AI Machines",
SIGART pp.21-28, (July 1980).

11-Apr-81(Sat)13-15 Akiyama (Enter Matsubara)
Masterson, "A separation-Individuation Failure:Interpersonal",
chapter 3 of Psychotherapy of the borderline adult.
"Ishiki-no-shougai"

25-Apr-81(Sat)14-16 Ogasawara
Winograd, "What does it mean to understand language?",
Cognitive Science 4, pp.209-241

9-May-81(Sat)14-16 Katagiri?
?, Knowledge representation?

23-May-81(Sat)14-16 Saito
S.Turkle, "Computer as rorschach: Subjectivity and social responsibility",
MIT memo.

6-Jun-81(Sat)13-15 Shirai
H.Levesque & J.Mylopoulos, "a procedural semantics for semantic networks",
pp.93-120, Associative Networks(Findler ed.).

20-Jun-81(Sat)14-16 Suzuki (Enter Hori)
Alan Bundy & Bob Welham, "Using Meta-level Inference for Selective
Application of Multiple Rewrite Rule Sets in Algebraic Manipulation",
AI vol.16, pp.189-212, (1981).
Alan Bundy & Bob Welham, "Using Meta-level Description for Selective
application of Multiple Rewrite Rules in kAlgebraic Manipulation",
Working Paper 55, Dep. of AI Edinburgh Univ. (May 1979).

11-Jul-81(Sat)14-16 Nakashima
Hector J. Levesque, "A procedural Approach to Semantic Networks",
Technical Report No.105, Dep. of CS, Univ. of Toronto (April 1977).

18-Jul-81(Sat)14-16 Nakashima
ditto

1-Aug-81(Sat)..3-Aug-81(Mon) AI summer seminar
Hector J. Levesque, "A procedural Approach to Semantic Networks",
Technical Report No.105, Dep. of CS, Univ. of Toronto (April 1977).

12-Sep-81(Sat)14-16 Mizoguchi
Fernando C.N. Pereira and David H. D. Warren,
"Definite Clause Grammars for Language Analysis - A Survey of the
Formalism and a Comparison with Augmented Transition Networks",
AI vol.13, pp.231-278, (1980).

26-Sep-81(Sat)14-16 Mizoguchi
ditto

17-Oct-81(Sat)14-16 Katagiri
Robert W. Lawler, "The Progressive Construction of Mind",
Cognitive Science vol.5, pp.1-30, (1981).

7-Nov-81(Sat)14-16 Ogasawara
Richard E. Fikes, "Odyssey: A Knowledge-Based Assistant",
AI vol.16, no.3, pp.331-361, (1981).

(28-Nov-81(Sat)15-17 Kumon did not show up)

5-Dec-(Sat)15-17 Kumon
E.Charniak, "A Common Representation for Problem-solving and
Language-Comprehension Information",
AI vol.16, pp.225-255, (1981).

19-Dec-81(Sat)14-16 Saito
P.N. Johnson-Laird, "Mental Models in Cognitive Science",
Cognitive Science vol.4, pp.71-115, (1980).

1982

9-Jan-82(Sat)15-17 Shirai
Roger C. Schank & Christopher K. Riesbeck,
"Inside Computer Understanding: Five Programs plus Miniatures",
Chapters 5-8, Lawrence Erlbaum Associates, Publishers (1981).

23-Jan-82(Sat)17-19 Suzuki
Larry Tesler, "The Smalltalk Environment", pp.90-147
James C. Althoff Jr.,
"Building Data Structures in the Smalltalk-80 System", pp.230-278
L Peter Deutsch,
"Building Control Structures in the Smalltalk-80 System", pp.322-346
BYTE vol.6, no.8, special issue on Smalltalk (August 1981).

6-Feb-82(Sat)13-15 Saito, Katagiri
Alan Borning, "The Programming Language Aspects of ThingLab,
a Constraint-Oriented Simulation Laboratory",
ACM Trans. on P. L. and Systems, vol.3, no.4, pp.353-387, (Oct 1981).

13-Feb-82(Sat)13-15 Nakashima
ditto

27-Feb-82(Sat)13-15 Hori
Albus, "Brains, Behavior, and Robotics",
BYTE books.

20-Mar-82(Sat)13-15 Nakashima
Ehud Y. Shapiro, "Inductive Inference of Theories From Facts",
Research Report 192, (February 1981).

10-Apr-82(Sat)13-15 Nakashima
ditto

17-Apr-82(Sat)13-15 Horimoto
Jurgen M. Janas & Camilla B. Schwind,
"Extensional Semantic Networks: Their Representation, Application, and Generation",
In Associative Networks (N. V. Findler ed.), pp.267-302, (1979).

(8-May-82(Sat)13-15 Matsukata postponed)
22-May-82(Sat)13-15 Matsubara
Paul S. Rosenbloom, "A World-Championship-level Othello Program",
CMU-CS-81-137, (August 1981).
1981 July issue of BYTE (related article).
5-Jun-82(Sat)13-15 Matsukata
Michael R. Genesereth, Russell Greiner, and David Smith, "MRS Manual",
Stanford HPP Memo HPP-80-24, (December 1980).
19-Jun-82(Sat)13-15 Mizoguchi
Mark Stefik et. al., "The Organization of Expert Systems, A Tutorial",
AI vol.18, pp.135-173, (1982).
3-Jul-82(Sat)13-15 Aida
Allen Newell, "The Knowledge Level",
AI vol.18, pp87-127, (1982).
17-Jul-82(Sat)13-15 Katagiri
Wendy G. Lehnert, "Plot Units and Narrative Summarization",
Cognitive Science vol.4, pp.293-331, (1981).

30-Jul-82(Fri)..1-Aug-82(Sun) AI summer seminar
Karl R. Popper & John C. Eccles,
"The Self and Its Brain -- an argument for Interactionism",
Springer International (1981).
Part I
P1 Materialism Transcends Itself Aida
P2 The Worlds 1,2 and 3 Katagiri
P3 Materialism Criticized Nakashima
P4 Some Remarks on the Self Goto
P5 Historical Comments on the Mind-Body Problem Saito
P6 Summary
Part II
E1 The Cerebral Cortex
E2 Conscious Perception
E3 Voluntary Movement
E4 The Language Centres of the Human Brain
E5 Global Lesions of the Human Cerebrum
E6 Circumscribed Cerebral Lesions
E7 The Self-Conscious Mind and the Brain Mizoguchi
E8 Conscious Memory:
The Cerebral Processes Concerned in Storage and Retrieval Matsubara
Part III
Dialogue I Sato
Dialogue II Shimada
Dialogue III Shirai
Dialogue IV Suzuki
Dialogue V Tada
Dialogue VI Terano
Dialogue VII Tomura
Dialogue VIII Nakagawa
Dialogue IX Kumon
Dialogue X Hori
Dialogue XI Matsukata
Dialogue XII Suzuki

21-Aug-(Sat)9-17
Karl R. Popper & John C. Eccles,
"The Self and Its Brain -- an argument for Interactionism",
Springer International (1981).
11-Sep-82(Sat)13-15 Kumon (Constraints start)
Judea Pearl, antonio Leal, and Joseph Saleh,
"GODDESS: A Goal-Directed Decision Structuring System",
IEEE Trans. on PAMI vol.PAMI-4, no.3, pp.250-262, (May 1982).

Constraints separately going

Guy Lewis Steele Jr.,
"The definition and implementation of computer programming language
based on Constraints",
MIT doctoral dessertation.
11-Sep-82(Sat)16-18 Shirai
25-Sep-82(Sat)16-18 Nakagawa?
9-Oct-82(Sat)16-18 Katagiri
30-Oct-82(Sat)16-18

Members as of 3-Jul-1982

Aida, Akiyama, Ogasawara, Katagiri, Kishi, Kumon, (Goto), Saito, Sato, Shimada,
Shirai, Suzuki, Tada, Terano, (Tomura), Nakagawa, Nakashima, Hori, Horimoto,
Matsukata, Matsubara, Mizoguchi, (Yokoo)

25-Sep-82(Sat)14-16 Saito
Patrick J. Hayes, "The Naive Physics Manifesto",
In Expert systems in the micro electronic age (D.Michie ed.), pp.242-287, (1979).
9-Oct-82(Sat)13-15 Shimada
R.Wilensky, "Meta-Planning: Representing and Using Knowledge About Planning
in Problem Solving and Natural Language Understanding",
Cognitive Science vol.5 ,pp.197-233, (1980).
30-Oct-82(Sat)13-15 Shirai
Reid G. Smith,
"A Framework for Distributed Problem Solving", chapters 1-3,
UMI research Press, Computer Science:Artificial Intelligence, No.10.
6-Nov-82(Sat)10-17 Saito, Mizoguchi, Matsubara
Karl R. Popper & John C. Eccles,
"The Self and Its Brain -- an argument for Interactionism",
Springer International (1981).
13-Nov-82(Sat)13-15 Terano
Reid G. Smith,
"A Framework for Distributed Problem Solving", chapters 4-7,
UMI research Press, Computer Science:Artificial Intelligence, No.10.
4-Dec-82(Sat)13-15 Suzuki
Giuseppe Attardi & Maria Simi, "Extending the Power of Programming by Examples",
Proc. of SIGOA Conference on Office Information Systems, pp.52-66, (June 1982).
18-Dec-82(Sat)13-15 Nakagawa
M.P.Georgeff, "Procedural Control in Production Systems",
AI vol.18, pp.175-201, (1982).

1983

8-Jan-83(Sat)15-17 Nakashima
Hideyuki Nakashima, "A Knowledge Representation System: Prolog/KR",
Doctral Thesis submitted to the University of Tokyo (1982).
22-Jan-83(Sat)14-16 Nakashima
ditto
29-Jan-83(Sat)14-16 Nakashima
ditto
5-Feb-83(Sat)14-16 Hori
P.H.Winston, "Learning New Principles From Precedents and Exercises",
AI vol.19, pp.321-350, (1982).
26-Feb-83(Sat)13-15 Hashida
Koiti Hashida, "A Real-time Syntacitc Analysis of Natural Languages",
M.Sc Thesis submitted to the University of Tokyo (Jan 1983).

19-Mar-83(Sat)13-15 Hashida
ditto
2-Apr-83(Sat)15-17 Hiraga
Drew McDermott, "A Temporal Logic for Reasoning About Processes and Plans",
Cognitive Science, vol.6, pp.101-155, (1982).
16-Apr-83(Sat)15-17 Hiraga
ditto
7-May-83(Sat)15-17 Matsukata
Ryszard S. Michalski, "A Theory and Methodology of Inductive Learning",
AI vol.20, pp.111-161, (1983).
21-May-83(Sat)15-17 Matsukata
ditto
28-May-83(Sat)15-17 Matsubara
Gerald Jay Sussman and Guy Lewis Steele Jr., "CONSTRAINTS -- A Language
for Expressing Almost-Hierarchical Descriptions",
AI vol.14, pp.1-39, (1980).
11-Jun-83(Sat)15-17 Aida
Lewis M. Norton, "Automated Analysis of Instructional Text",
AI vo.20, pp.307-344, (1983).
25-Jun-83(Sat)15-17 Katagiri
Jean Mark Gawron et al.,
"Processing English with a Generalized Phrase Structure Grammar",
ACL Annual Conference, pp.74-81, (1982).
(Association for Computational Linguistics)
9-Jul-83(Sat)15-17 Tanaka
Takushi Tanaka, "Representation and analysis of electrical circuits
in a deductive system",
Proc. IJCAI, Karlsruhe, W.Germany. (1983).

18-Aug-83(Thurs)..21-Aug-83(Sun) AI Summer Seminar
Douglas R. Hofstadter, Daniel C. Dennett,
"The MIND'S I Fantasies and Reflections on Self and Soul",
Basic Books Inc. (1981).
I. A Sense of Self Hori, Kishi, Kumon
II. Soul Searching Shirai, Nakashima, Hiraga
III. From Hardware to Software Hashida, Yokoi, Aida
IV. Mind as Program Matsubara, Matsukata, Terano
V. Created Selves and Free Will Saito, Katagiri, Nakagawa
VI. The Inner Eye Shimada, Shinohara, Tada

10-Sep-83(Sat)15-17 Kishi, Shirai
Robert C. Berwick, "Introduction: Computational Aspects of Discourse",
Chapter 1 of ?
17-Sep-83(Sat)14-16 Shimada, Shinohara, Tada
Douglas R. Hofstadter, Daniel C. Dennett, "The MIND'S I",
Chapter VI. The Inner Eye, Basic Books Inc. (1981).
1-Oct-83(Sat)15-17 ?
?
8-Oct-83(Sat)15-17 Kishi, Shirai
Robert C. Berwick, "Introduction: Computational Aspects of Discourse",
Chapter 1 of ?
22-Oct-83(Sat)15-17 Kishi, Shirai
ditto
5-Nov-83(Sat)15-17 Saito
Ken Forbus, "Qualitative Process Theory",
MIT AI Lab Memo No.664, (Feb 1982, Revised June 1983).
19-Nov-83(Sat)15-17 Shinohara
Marvin Minsky, "Jokes and the Logic of the Cognitive Unconscious",
MIT AI Memo, No.603, (1980).
17-Dec-83(Sat)15-17 Shimada
Douglas B. Lenat, "The Nature of Heuristics",
AI vol.19, pp.189-249, (1982).
"Theory Formation by Heuristic Search
The Nature of Heuristics II: Background and Examples",
AI vol.21, pp.31-59, (1983).
"EURISKO: A Program That Learns New Heuristics and Domain Concepts
The Nature of Heuristics III: Program Design and Results",
AI vol.21, pp.61-98, (1983).

1984

7-Jan-84(Sat)15-17 Tanaka
Zohar Manna and Amir Pnueli,
"Verification of Concurrent Programs, Part I:The Temporal Framework",
Research Report No.STAN-CS-81-836, (June 1981).
28-Jan-84(Sat)15-17 Nakagawa, Terano
Jon Doyle, "A Society of Mind
Multiple perspectives, reasoned assumptions, and virtural copies",
8th IJCAI, pp.309-314, (1983).
18-Feb-84(Sat)15-17 Nakashima
Yorick Wilks and Janusz Bien,
"Beliefs, Points of View, and Multiple Environments",
Cognitive Science, vol.7, pp.95-119, (1983).
25-Feb-84(Sat)15-17 Terano
?, "XPLAIN",
AI, vol.21, no.3, pp.285-?, (?).
10-Mar-84(Sat)15-17 Hiraga
Derek Bickerton, "Creole Languages",
Scientific American, vol.249, no.1, (July 1983).
31-Mar-84(Sat)15-17 Hashida
Dedre Gentner, "Structure-Mapping: A Theoretical Framework for Analogy",
Cognitive Science, vol.7, pp.155-170, (1983).

6-Apr-84..8-Apr-84 AI Spring Seminar
J. Fodor, "Modularity of Mind"

21-Apr-84(Sat)15-17 Hori
Patrick H. Winston, Thomas O. Binford, Boris Katz and Michael Lowry,
"Learning Physical Description from Functional Definitions,
Examples and Precedents",
Stanford Report No.STAN-CS-82-950 or MIT AI Memo No.AIM-349
12-May-84(Sat)15-17 Aida
Joxan Jaffar, Jean-Louis Lassez and John Lloyd,
"Completeness of the Negation as Failure Rule",
Proc. of the 8th IJCAI, pp.500-506, (1983).

26-May-84(Sat)15-17 Matsubara
Josef Psotka,
"Perceptual Processes That May Create Stick Figures and Balance",
Journal of Experimental Psychology, Human Perception and Performance,
vol.4, no.1, pp.101-111, (1978).
16-June-84(Sat)15-17 Kishi
Paul Lefrere, "Text Processing",
chapter 13 from Artificial Intelligence - Tools, Techniques,
and Applications by Tim O'Shea and Marc Eisenstadt, Harper & Row.
7-July-84(Sat)15-17 Koyama
William A. Lornfeld, "Equality for Prolog",
Proc. of the 8th IJCAI, pp.514-519, (1983).
21-July-84(Sat)15-17 Saito
?