The CREDO stack: theory to practice in cognitive systems design

John Fox[1] JOHN.FOX@ENG.OX.AC.UK
Lincoln College, University of Oxford

Abstract

CREDO is a long term research program in cognitive science whose aim is to develop a theoretical framework for understanding general cognition as a basis for designing cognitive systems, and validating the approach in practical applications (primarily in medicine). This short overview summarises results of six main sub-programs, from empirical studies to theory to practical technologies and applications. The ideas at each level are illustrated using a running example of a cognitive agent designed to support clinicians in the care of patients with suspected or proven COVID-19 infections (accessible at https://www.openclinical.net/index.php?id=746). The CREDO model builder can be downloaded from OpenClinical.net, in order to design, build and test CREDO style cognitive systems and for the community to critically assess the CREDO methodology and its foundations.

1. Introduction

CREDO is a framework for understanding and designing cognitive systems. It has evolved through a multi-decade program of empirical studies of clinical expertise and the development of practical technologies and applications. Medicine has long been fertile soil for research in cognitive science and systems engineering because of its complexity and the range of challenges that it throws up for theoreticians as well as practical developers. It is particularly challenging because of the diversity of clinical tasks; the high levels of uncertainty that are ubiquitous in clinical practice; the need for clinicians to address multiple goals and constraints concurrently and adapt to changing and unexpected situations. A major challenge is the vast amount of background knowledge that may be relevant at any time in clinical practice.

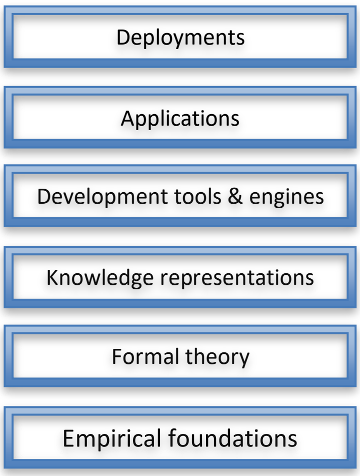

Figure 1: The CREDO stack is a hierarchical view of CREDO concepts and techniques we have developed for modeling cognition in complex domains and achieving/exceeding human level performance on routine but non-trivial tasks

Medicine has been a particular focus for research in psychology and AI, decision theory (descriptive and normative), mathematics (e.g. Bayesian inference), computer science (e.g. qualitative reasoning and formal logic), and knowledge engineering (e.g. expert systems, ontology modelling). These perspectives have come together in a design framework summarized as the CREDO stack (figure 1). Although we have taken medicine as a specific research domain the goal has been to create a general theoretical foundation for understanding high level cognition and practical methods for developing general purpose cognitive agents (Das et al, 1997; Fox and Das, 2000; Fox et al, 2006, 2013).

1.1 Summary position

Starting from empirical studies of medical decision making and clinical practice the CREDO program has evolved a benchmark set of cognitive capabilities which an autonomous agent would need to carry out a

[1] As a newcomer to the ACS community I thank the three reviewers whose constructive criticism and suggestions have led to significant improvements to this paper.

large proportion of routine medical tasks (Fox et al, 2003). It is possible to develop a formal model of these capabilities as a set of abstract “signatures” which can be instantiated in the diverse ways which are preferred in different fields of cognitive science (e.g. mathematical modelling; cognitive simulation; system architectures) to promote interdisciplinary collaborations (Fox et al 2013). While inspired by human expertise CREDO is not primarily a theory of human cognition but has been shown to be an effective foundation for developing cognitive agents that clinicians find natural and can match or outperform clinician performance on non-trivial medical tasks.

1.2 Running example – support for cognitive tasks in the care of COVID-19 patients

We have recently completed the first phase of a project that has used the CREDO framework to create, validate and publish a computational model of best clinical practice in the detection, diagnosis and management of patients with COVID-19 infections**.** The purpose of the “Pathfinder” system is to (1) demonstrate a cognitive approach to the design of AI services for assessment by clinicians and researchers, healthcare providers, technology developers and others seeking to address the pandemic emergency, (2) develop point of care products and services which embody best clinical practice in data interpretation, decision-making, workflow management and other “intelligent” services across the patient journey “from home to hospital to home”. Pathfinder is used here to illustrate features of CREDO at each level of the stack**.**

2. History and overview

2.1 Empirical foundations

The origins of the CREDO program were in studies of human decision-making and computer simulations of cognitive processes involved, notably memory, reasoning, decision-making and learning (e.g. Fox, 1980; Cooper and Fox, 1997; Cooper and Yule, 1997). These studies suggested that human decision-making is at least as well described by a symbolic account of complex reasoning as by traditional decision models. Our early work used standard production rule techniques to simulate knowledge based processes (e.g. Fox 1980, Fox et al 1980) but as many in the medical world have noted rules have significant limitations in practical use (e.g. Musen, 1998).

Working closely with clinicians over many years has led to the view that a single form of inference does not reflect the diverse features of clinical reasoning and decision making. To address this we developed a framework which distinguishes a number of distinct types of reasoning.

1. Inference over *beliefs* (e.g. about a clinical situation)
2. Raising *goals* in response to beliefs (e.g. diagnose D, assess risk of D, test for D, treat D)
3. Formulating alternative candidate solutions for achieving goals (*decision options)*;
4. Construct *arguments* for and against each decision option;
5. *Aggregate* arguments for and against competing options to give a preference order over the alternatives;
6. *Commit* to one or more of the options

Unlike classical decision this account was initially informal and descriptive, not normative or axiomatic.

Pathfinder *demonstrates reasoning about situations yielding beliefs about possible diagnoses (e.g. COVID-19, common pneumonia and a dozen other candidates) and tasks to achieve goals (e.g triage, select tests, predict complications and prognosis). Pathfinder constructs arguments for and against alternative decision candidates, some of which are aggregated quantitatively (when statistical data are available) or with more human-like semi-quantitative methods (pros and cons). The Pathfinder decision agent does not commit to decisions autonomously (though technically it can) but makes recommendations for human approval.*

2.2 Cognitive theory

A key question about traditional cognitive systems is how to ensure they are trustworthy since, despite their practical successes, the *ad hoc* nature of rule-based and other symbolic techniques compares unfavorably with formal mathematical models. However, while classical decision systems claim a "rational" (or at least optimal) decision procedure they show little of the versatility and flexibility of human cognition. For example they say little about when decisions should be initiated and what the options are, which domain-specific knowledge is relevant, or how to explain the rationale for recommendations (the famous "black box" issue). An early goal therefore was to develop a theoretical framework covering the whole decision cycle. This was dubbed *Symbolic Decision Theory* (Fox et al, 1990a; Fox and Krause 1991; Huang et al, 1993).

Another driver for developing formal theoretical foundations for the CREDO framework was the desire to understand decision-making, planning and other tasks by autonomous agents, a traditional research focus in AI, particularly for safety critical applications such as clinical decision making. The domino agent model shown in figure 2 extended the scope of symbolic decision theory to include planning and plan execution (figure 2; Das et al, 1997) and grounded the informal theory in 2.1 in a well defined semantics (Fox and Das, 2000)

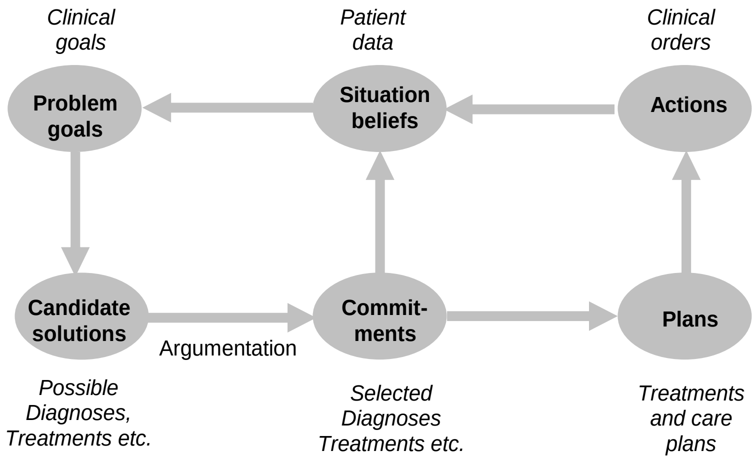

Figure 2: the domino agent model brings together concepts of beliefs, goals, reasoning, decision making, planning and acting within a unified framework of non-classical logics (Das et al 1997; Fox and Das, 2000).

Pathfinder*: the on-line demonstration incorporates most of the cognitive functions in the domino, including continuous updating of beliefs, argumentation and recommendations for most decisions in the COVID domain (triage diagnosis, prognosis, prediction of complications and selection of treatment plan). In addition the agent can manage the workflow required in a treatment or other care plan, and update the plan should circumstances change. The agent can explain many aspects of its decision making a current area of weakness is in explaining plans, such as the history of current actions and intended future plans and their rationale*.

2.3 The PRO*forma* language

The domino model is the basis of an agent programming language called R_2L whose semantics are formalized in classical and non-classical logics (Das et al, 1997; Fox and Das, 2000). However a more practical need is to support rigorous design and engineering of cognitive systems drawing on good engineering practice (e.g. system modularity, component reusability). R_2L consequently evolved into the PRO*forma* language which combines strengths of many computing paradigms in an intuitive agent modelling language (Sutton and Fox, 2003).

The key idea in PRO*forma* is that the network of *logic processes* in the domino model (arrows) was reified into a small set of *semantic objects* representing a small ontology of cognitive tasks: decisions, plans, actions and enquiries (figure 3). These four ontological classes are subclasses of a notional abstract task called a “keystone” from which they inherit attributes such as the *goal* of the task, *pre-conditions*, *post-conditions* and *triggers*.

The four main task classes have class-specific attributes as follows

- *Decisions* have attributes which specify the candidates or *options* of a decision, logical rules for constructing arguments for and against each option, and criteria for choosing or recommending one or more of the options .
- *Plans* are networks of tasks, which are made up of component tasks (including sub-plans), sequencing constraints (if any); plans may also include conditions in which the plan should terminate (indicating success) or abort, possibly triggering other plans.
- *Actions* and *enquiries* are the simplest tasks whose role is to communicate with the agent’s external environment, as when interacting with a user or other software systems through a conventional UI or devices.

PRO*forma* task networks are easy for non-programmers such as clinicians to understand, many of whom have successfully developed executable models of particular clinical expertise. The. language syntax and execution semantics are standardized (Sutton and Fox, 2003) and several execution engines are available. It has proved to be a highly versatile language for modelling medical and other kinds of human expertise.

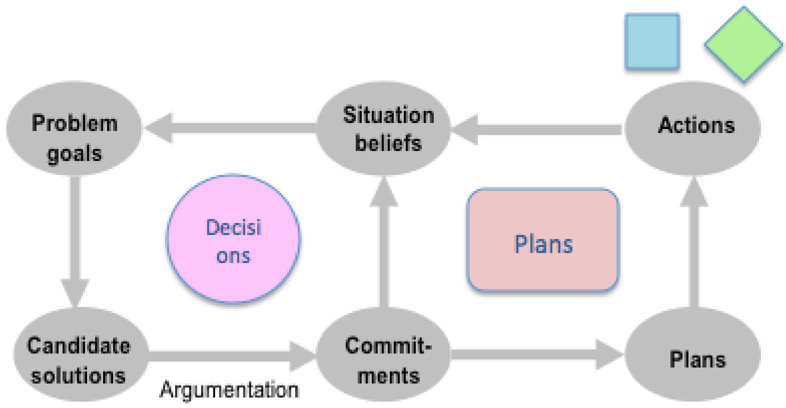

Figure 3: Reification of the domino model into “tasks”. Arrows (logics) on the left of the domino are reified into a generic decision task, and the arrows (logics) on the right are reified as plans.

***Pathfinder** is implemented entirely in PROforma. It represents knowledge of best practice in the care of people with COVID-19 infections as this was understood in March-April 2020. The model includes a variety of decisions and plans in a modular design, with some of the decisions and other “subskills” reused in different sub-plans.*

2.4 Applications

The earliest PRO*forma* applications were developed to support physicians in making routine decisions, including diagnosis, risk assessments, drug prescribing etc. Later developments focused on specialist care, particularly cancer care such as treatment of leukemia in children, diagnosis and treatment of breast cancer, genetic risk management and other complex domains. An important class of applications supports decision-making by multidisciplinary clinical teams (e.g. Patkar et al, 2006, 2011) who use it to support their decision making, treatment planning and patient monitoring in different settings on multiple occasions Trials of PRO*forma* applications have demonstrated benefits in better clinical decision making, and in some cases dramatic improvements over unaided clinician’s judgments. A recent review identified 16 trials showing performance superior to that of clinicians which have been published in peer reviewed medical journals (Fox et al; 2020)

Pathfinder *has not been trialed in clinical use and, given the speed at which knowledge of best practice in the management of COVID-19 infections is evolving, but a trial is planned. However, a commercial partner has used its own variant of the CREDO framework to build and deploy a COVID-19 patient assessment application that is in routine clinical use in a large UK hospital (see next section).*

2.5 Deployments

The top layer of the CREDO stack is a collection of services for deploying PRO*forma* applications, such as user interface services. The primary purpose of the language was to formalize decisions, plans and other tasks in a machine interpretable format, but since the CREDO task ontology and argument-based decision making are inspired by human cognitive processes we hoped this would facilitate intuitive user interfaces and interactions.

The simplest arrangement is to map each PRO*forma* task that is being executed to a particular style of web page. For example a *decision* is visualized as a single page with a set of options, each of which can be expanded to show the arguments for and against each option, and each argument can be selected to review the logical structure and backing evidence for it. An *enquiry* maps to a page that provides an electronic form for data entry, and an *action* might map to a page that presents an alert or a message to the user using a variety of widget types. Finally, a *plan* can be visualized as a collection of tasks, where a workflow is visualized as a network of enquiries, decisions, plans and actions organized as an executable network (see https://www.openclinical.net/index.php?id=68).

The presentation style of any task UI can be customized but such customizations are not considered part of the cognitive design theory or CREDO methodology. However, some studies have shown that the task ontology and process modeling framework are an effective basis for integrating perceptual modules (e.g. medical image processing; Taylor et al, 1999; Sordo et al 2001) and spoken language interfaces (Beveridge and Fox, 2006). The latter demonstrates a flexible, mixed initiative dialogue (see https://www.openclinical.net/index.php?id=371&searched=diva&advsearch=oneword&highlight=ajaxSearch_highlight+ajaxSearch_highlight1).

The Pathfinder UI *is "out of the box" as described above with minor customisation. However the patient assessment system running at Liverpool University hospital mentioned above, which also uses PROforma for knowledge modelling, has a highly customized UI which is optimized for use in the demanding setting of the emergency department of a busy hospital (www.deontics.com/news).*

3. Discussion

3.1 Comparison with other work

The concept of a *task* (keystones, decisions, plans etc) is central to the CREDO knowledge representation and the PRO*forma* language. As in *Cognitive Task Analysis* (Crandall et al, 2006) a task is a key element in design, but in the CREDO framework a task is a process that is intended to achieve a goal, and a PRO*forma* task is an executable model for doing that. Each task may be a simple action or a complex network of decisions, actions, plans and sub-plans; each PRO*forma* task is an object with a set of general attributes inherited from the abstract "keystone" task, augmented with a task specific set of attributes populated with domain-specific content.

In a previous surveyof the CREDO program (Fox, 2017) I distinguished between four theoretical traditions in cognitive science: *statics* (e.g. information processing architectures); *dynamics* (e.g. rule-based cognitive modelling); *epistemics* (e.g. knowledge representation) and "*anthropics*" (e.g. the mental states of folk psychology and philosophy of mind). We now briefly compare the CREDO approach with these traditions.

Statics: "Information processing architectures" and "cognitive architectures" are well known (e.g. ACT-R, CLARION, EPIC, SOAR). This is not a focus of the CREDO program which emphasizes cognitive *function* (e.g. reasoning, decision making, planning) rather than *structure* (box-arrow models). In the CREDO view key functions can be implemented on many different architectures: some may be human-like (e.g ACT-R) and some not (e.g. subsumption and 3-layer robot architectures). It is of interest however that as with ACT-R the key elements of the domino model can be mapped to a neuropsychological theory of the organization of human frontal lobe (Shallice and Cooper, 2011, chapter 9).

Dynamics: Modelling cognition using rule-based systems (e.g. OPS5, SOAR, ACT-R) are seen as implementation choices rather than fundamental to computational theories of cognition. However we have discussed elsewhere the notion of a "canonical theory" of cognitive processes which describes a framework for designing cognitive agents using generic function *signatures* that can be implemented in different ways such as production systems *but also* with *ad hoc* software engineering; logic programming; agent programming and other paradigms.

Epistemics: a lesson from knowledge representation research in AI is that general knowledge and domain-specific knowledge are not just an undifferentiated "soup" of concepts, rules etc. Knowledge is highly structured: it can be formalized at many levels of abstraction. The CREDO program has evolved a particular framework for modelling medical expertise that we call the knowledge ladder. This is described in the COVID-19 paper (*op cit*) which demonstrates how different representations can support distinct cognitive capabilities within a single application.

Anthropics: The domino model (Das et al, 1997) has similarities with the beliefs-desires-intentions framework originating in the philosophy of mind (Bratman, 1999) which is now popular in AI and agent systems; as events occur and situations change a domino agent can reason about its current situation (beliefs) with respect to its desires (goals) and intentions (plans), but it can also generate multiple (tentative) beliefs, goals and plans, argue the merits of the alternatives, and decide which to commit to in light of what it believes and knows.

3.2 Future work

The CREDO program has sought to develop a principled framework for designing, developing and deploying a wide range of cognitive systems which can carry out complex tasks at or above the level of human experts. The challenge domain has been medicine, but we think the approach and tools are potentially applicable to many other domains of human expertise. Medicine continues to be a rich source of scientific questions and engineering ideas, but we hope to validate the general framework in other non-medical domains.

The 2003 benchmark for a general "cognition engine" will need to be refined and extended in significant ways. Emerging fields of AI, such as cognitive vision, cognitive robotics, natural language or multi-agent systems. are not our areas of primary expertise so we are interested to explore collaborations with others in the ACS community on wider theoretical and applied questions. Our urgent objective is to update our 2003 benchmark and develop a new language with broader scope and software tools that can support the development and deployment of advanced cognitive systems. We are currently developing the requirements for the new language (working title PRO*forma* 2020) and would welcome comment and criticism of this work.

4. Conclusions

CREDO is a productive framework for understanding human expertise and engineering cognitive systems. The PRO*forma* language for specifying and implementing cognitive agents has proved to be very versatile and is an important element of many systems that are being routinely used in a variety of clinical settings. The development tools are constantly evolving to improve the versatility and flexibility of this approach to cognitive systems engineering but the basic cognitive principles have remained remarkably stable since first published in 1996 and the language was standardized in 2003.

References

Beveridge M, Fox J "Automatic generation of spoken dialogue from medical plans and ontologies" *Journal of Biomedical Informatics,* Volume 39 Issue 5, October 2006 Pages 482 – 499

Bratman M E *Intentions, plans and practical reason* Stanford University: CSLI, 1999

Card S, Moran P, Newell A *The psychology of human-computer interaction.* Lawrence Erlbaum Associates (1983).

Cooper R, Fox J Learning to make decisions under uncertainty: The contribution of qualitative reasoning *Proceedings of the 19th Annual Conference of the Cognitive Science Society*, 1997

Cooper, R. P., Yule, P., and Fox, J. (2003). Cue selection and category learning: a systematic comparison of three theories. *Cogn. Sci. Quarterly* 3, 143–182.

Crandall B, Klein G, Hoffman R *Working Minds*, Cambridge: MIT 2006.

Das S, Fox J, Elsdon D, Hammond P "A flexible architecture for a general intelligent agent*", Journal of Experimental and Theoretical Artificial Intelligence*, 9, 407-440, 1997.

Das S "Symbolic Argumentation for Decision Making under Uncertainty" in *7th International Conference on Information Fusion (FUSION),* 2005

Elvang-Goransson M, Krause P and Fox J "Acceptability of arguments as logical uncertainty" in M Clarke, R Kruse and S Moral (eds) Symbolic and Quantitative approaches to reasoning and uncertainty *Proc. European Conference ECSQUARU 93*, Berlin: Springer Verlag, 79-84.

Fox, J Making decisions under the influence of memory. *Psychological Review,* Vol 87(2), Mar 1980, 190-211

Fox J, Barber D and Bardhan C Alternatives to Bayes? A quantitative comparison with rule-based diagnostic inference *Methods of information in Medicine*, Oct;19(4):210-5. 1980

Fox J, Krause P J, Symbolic decision theory and autonomous systems, Proceeding UAI'91 *Proceedings of the Seventh conference on Uncertainty in Artificial Intelligence* Pages 103-110 San Francisco: Morgan Kaufmann Inc. 1991)

Fox J, Krause P J and Elvang-Goransson M "Argumentation as a general framework for uncertain reasoning", *Proc. Int. Conf. on Uncertainty in Artificial Intelligence,* Washington, 1993

Fox J and Das S, *Safe and Sound: Artificial Intelligence in Hazardous Applications*, AAAI/MIT Press, 2000

Fox J, Beveridge M, Glasspool D "Understanding intelligent agents: analysis and synthesis" *AI Communications* Volume 16 Issue 3, Pages 139 - 152, August 2003

Fox J, Glasspool D, Modgil S "A canonical agent model for healthcare applications" *IEEE Intelligent Systems: Special Issue on Intelligent Agents in Healthcare*, 21-28, Nov/Dec 2006.

Fox J, Cooper R, Glasspool D W "A canonical theory of dynamic decision-making" *Frontiers in Psychology*, 02 April 2013

Fox et al OpenClinical.net: AI and knowledge engineering at the point of care *British Medical Journal, health and care informatics,* 2020 (in press)

Krause, P., Ambler, S., Elvang-Goransson, M. and Fox, J. (1995) A logic of argumentation for reasoning under uncertainty, *Computational Intelligence*, 11, 113{131

Musen M "Modern Architectures for Intelligent Systems: Reusable Ontologies and Problem-Solving Methods" *Proceedings of AMIA*, February 1998.

Patkar, V Acosta D …, Keshtgar M Using computerised decision support to improve compliance of cancer multidisciplinary meetings with evidence-based guidance *BMJ Open*. 2012; 2(3):2012

Patkar V et al Evidence-based guidelines and decision support services: A discussion and evaluation in triple assessment of suspected breast cancer British Journal of Cancer 95(11):1490-6

Taylor P et al The development and evaluation of CADMIUM: a prototype system to assist in the interpretation of mammograms *Medical Image Analysi*s, 1999, 3(4), 3231-337.

Shallice T and Cooper R *The organization of mind* Oxford: Oxford University Press, 2011.

Sordo M et al Combining Decision Support and Image Processing: A PROforma Model V. Patel et al. (Eds) MEDINFO 2001 Amsterdam: IOS Press

Sutton D and Fox J, "Syntax and semantics of the PRO*forma* guideline modeling language", *JAMIA* 2003