Rethinking AI Strategy and Policy as Entangled Super Wicked Problems

Ross Gruetzemacher
Auburn University, Harbert College of Business, Department of Systems and Technology
rossg@auburn.edu

Abstract

This paper attempts a preliminary analysis of the general approach to AI strategy/policy research through the lens of wicked problems literature. Wicked problems are a class of social policy problems for which traditional methods of resolution fail. Super wicked problems refer to even more complex social policy problems, *e.g.* climate change. We first propose a hierarchy of three classes of AI strategy/policy problems, all wicked or super wicked problems. We next identify three independent super wicked problems in AI strategy/policy and propose that the most significant of these challenges – the development of safe and beneficial artificial general intelligence – to be significantly more complex and nuanced, thus posing a new degree of 'wickedness.' We then explore analysis and techniques for addressing wicked problems and super wicked problems. This leads to a discussion of the implications of these ideas on the problems of AI strategy/policy.

Introduction

In recent years, the scope and complexity of the challenges posed by the rapid proliferation of AI technology have begun to be recognized as a significant problem. Strategy and policy research has begun, but due to the complexity of the problem and uncertainties regarding risks, timelines and more broadly, the formulation of the problem, progress has been difficult. Similar policy problems are commonly considered to be wicked problems.

Wicked problems refer to a class of social policy problems for which scientific bases for resolution are insufficient and inappropriate (Rittel and Webber 1973). For such ill-defined problems, no definitive solution exists and *optimal solutions* can be formulated only with severe qualifications. Wicked problems include policy problems such as healthcare or tax reform as well as broader social problems like terrorism, poverty and hunger. We further distinguish more complex wicked problems, *e.g.* climate change, to be super wicked problems (Levin, Cashore et al. 2007, Lazarus 2008).

AI strategy and policy poses a complex set of research questions that are numerous and varied. Generally, the development of AI strategy and policy regards not a single problem but a number of independent yet intertwined problems of substantial perplexity. We distinguish these problems as being either one of narrow AI strategy/policy or that of artificial general intelligence (AGI) strategy/policy. A third class of AI strategy/policy problems also exists that we shall define as ancillary strategy/policy problems. We propose a hierarchy for consideration through which the nature of AI strategy/policy as a wicked problem may be best understood. As depicted in figure 1, at the top of the hierarchy is AGI strategy/policy followed by narrow AI strategy/policy and ultimately ancillary strategy/policy. The structure of this hierarchy is indicative of the degree of entanglement with the AGI strategy/policy problem being the most convoluted.

Narrow AI strategy/policy concerns complex issues such as mass job automation or the militarization of AI, which, for the purposes of this essay, shall be used to exemplify this class. In their own right, problems such as these constitute super wicked problems. This is to say that they represent planning problems of difficulty comparable to strategy/policy problems like climate change. This is true in so much as the fact that these problems embody the features of super wicked problems as described by Levin *et al*. (Levin, Cashore et al. 2007) .

The study of AGI strategy/policy[1] remains in its nascency. At present, significant strategy deficiencies are acting as a bottleneck for research involving implementable policy (Flynn 2017). Flynn describes the current state of the strategy problem as requiring significant 'disentanglement research' due to the high degree of entangled research questions. Due to the current nature of research on the AGI strategy problem, he suggests that those interested in AGI strategy/policy research not contribute directly at this time, rather accumulate skills and career capital until the bottleneck

[1] Due to the capacity of AGI for unprecedented societal transformation, we consider the problem of AGI strategy/policy to be the primary problem of concern. This is due to our consideration that strategy/policy relating to AGI regards the nature of policy concerning recursive, self-improving artificial general intelligence.

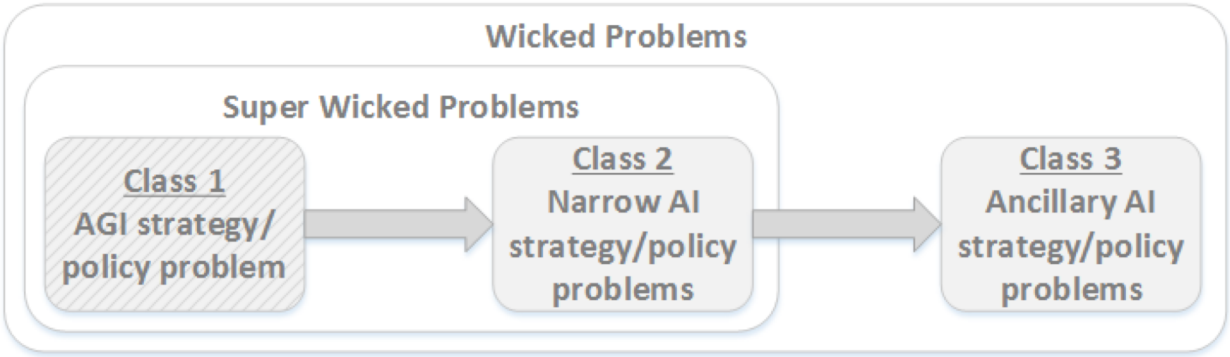

Figure 1: A hierarchical representation of the three classes of AI strategy/policy problems. The AGI strategy/policy problem is the sole class 1 problem. Class 2 problems are other entangled super wicked problems and class 3 problems are ancillary problems.

clears, presenting more clearly defined research questions that they may then pursue. We concur that the challenges posed by entangled research questions are significant, and we believe that the contents of this paper address such challenges by examining an alternative perspective through which to understand the problem.

In contrast to the former two classes, ancillary problems are more manageable – that is to say that problems of this class are merely traditional wicked problems. Examples include problems such as algorithmic accountability and the automation of healthcare systems. Problems of this class pose only marginal impacts on society and the future of humanity with respect to the potential impacts of problems of the first two classes. For this examination AGI and narrow AI strategy/policy problems are considered to be of critical concern, while ancillary problems, due to their relatively diminished impacts, are not, and will not be considered further.

This essay contributes to the body of existing literature in several significant ways. Firstly, this paper provides the sole attempt to understand the problem of planning for AI strategy/policy research through consideration of the problem as a wicked problem. Furthermore, the paper argues that AI strategy/policy planning cannot be considered a single wicked or super wicked problem, rather an entanglement of multiple super wicked problems of different hierarchical classes. Moreover, it suggests that these problems qualify the crucial problem of artificial general intelligence strategy/policy as a class of wicked problems unprecedented in complexity, uncertainty and fragmentation. Finally, it examines how the body of literature related to wicked problems and super wicked problems may be leveraged to benefit AI strategy/policy research.

Wicked Problems

Classic Wicked Problems

The term 'wicked problems' was first used over 50 years ago by Horst Rittel in describing a class of ill formulated, confusing social problems (Churchman 1967). The use of the adjective 'wicked' supposedly describes the mischievous and evil nature of these problems in defying attempts made to 'tame' them. In 1973 Rittel & Webber proposed wicked problems more formally as problems of 'social or policy planning' that are inherently incorrigible (Rittel and Webber 1973). They suggested that unlike well-posed problems of science and engineering, such problems could not be clearly defined and, consequently, had no verifiable solutions. They further described a wicked problem as never being solved, rather, at best, only being re-solved over and over again. They identified 10 characteristics of wicked problems which are presented in Table 1. These criteria are

10 Distinguishing Properties of Wicked Problems
1. There is no definitive formulation of a wicked problem.
2. Wicked problems have no stopping rule.
3. Solutions to wicked problems are not true-or-false, but good-or-bad.
4. There is no immediate test of a solution to a wicked problem.
5. Every solution to a wicked problem is a 'one-shot operation'; because there is no opportunity to learn by trial-and-error, every attempt counts significantly.
6. Wicked problems do not have an enumerable (or an exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan.
7. Every wicked problem is essentially unique.
8. Every wicked problem can be considered to be a symptom of another problem
9. The existence of a discrepancy representing a wicked problem can be explained in numerous ways. The choice of explanation determines the nature of the problem's resolution.
10. The planner has no right to be wrong.

Table 1: Rittel & Webber's 10 distinguishing properties of wicked planning-type problems.

not a set of tests, however, to objectively determine wickedness, rather, they are intended to offer insight for judging whether a problem is wicked (Camillus 2008).

An alternate formulation of wicked problems is presented by Head, who describes wicked problems as being a confluence of three factors: *complexity* of subsystems and interdependencies; *uncertainty* regarding risks and consequences of interventions; and *divergence* or fragmentation in values, viewpoints and strategic intentions (Head 2008). If problems were rated in each of these three dimensions, wicked problems would be those that score highly for each. This formulation of wicked problems is easily applied to problems arising from AI strategy/policy, as will be further examined in the discussion section.

Camillus suggests an alternate set of criteria for wicked problems, drawing on the 10 properties proposed by Rittel & Webber, but specific to problems of strategy in organizations (Camillus 2008). He describes five characteristics of wicked problems:

- The problem involves many stakeholders with different values and priorities.
- The issue's roots are complex and tangled.
- The problem is difficult to come to grips with and changes with every attempt to address it.
- The challenge has no precedent.
- There's nothing to indicate the right answer to the problem.

For addressing wicked strategy problems, he suggests increasing stakeholder involvement, documenting opinions and communication, *e.g.* brainstorming sessions and holding retreats for stakeholders. Similarly, Conklin perceives the difficulties associated with wicked problems as a consequence of project fragmentation (Conklin 2006). He proposes a dialogue mapping technique for building shared understandings among stakeholders to overcome this.

In the 1970s, Rittel & Webber were not alone in drawing critical attention to contemporary approaches for addressing complex policy problems. From a systems theory perspective, Ackoff described such problems as a *system of problems*, or more simply, as a *mess* (Ackoff 1974). Frequently, he explained, solutions could not be obtained by independently solving each of the constituent elements of the *mess*, suggesting that such efforts actually seemed to aggravate the situation. This notion of a *mess* is similar to the inherent entanglement of problems in wicked problems. In later work, Ackoff describes problem solving as the decisions made about four components: objectives, controlled variables, uncontrolled variables and the relationships among them (Ackoff 1978). He further describes it as having one of three successful outcomes: being solved, being resolved or being dissolved. Problems are solved when values of the controlled variables are identified which optimize the outcome. Problems are resolved when controlled variables are identified which result in an outcome that is good enough or satisficing. Problems are dissolved when altered such that the choices are no longer meaningful. Although not explicitly wicked problems research, parallel research like this can also be applicable.

Super Wicked Problems

The idea of a super wicked problem was first proposed by Levin *et al.* to describe the nature of the policy problem for global efforts to combat climate change (Levin, Cashore et al. 2007). That work set forth four central qualities of super wicked problems: 1) time is running out 2) there exists no central authority 3) those seeking to end the problem are also causing it 4) hyperbolic discounting. Extended descriptions of these four qualities can be found in the first column of Table 2. The remaining three columns of Table 2 are used to describe these qualities[2] in the context of the AGI strategy/policy problem and the two narrow AI strategy/policy problems.

Table 2 identifies the layers of super wicked problems comprising the AI broader strategy/policy problem. To dig deeper, we can consider the entangled nature of the two narrow AI strategy/policy problems and the AGI policy/strategy problem. For example, the militarization of AI[3] could lead to incredibly small and difficult to detect miniature drones being used for both passive and active acts of espionage in order to destabilize AI projects and steal sensitive materials. Alternately, mass unemployment from job automation in democratic societies could lead to the election of populist officials[4] (Frey, Berger et al. 2017) which could have numerous and unanticipated destabilizing effects[5].

We believe that the complexities and entanglement associated with multiple super wicked problems constitutes an increased degree of wickedness inherent in the problem, meriting consideration as a new class of wicked problem, *e.g.* a super, super wicked problem. The language, however, is insignificant, as may well be the distinction. Yet, the notion of the divergence of this problem from the existing class suggests that new theory, or modifications of the theory regarding the existing class, can be appropriate.

[2] These descriptions are not exhaustive.

[3] This example of the militarization of AI fails to illuminate the scope and complexity of potential negative outcomes. We note that the militarization of AI is not limited strictly to conventional warfare but can include autonomous and adaptive self-replicating digital agents as well as the development and application of other powerful cyberwarfare tools.

[4] Frey *et al*. suggest a link between job automation and the election of a populist in the 2016 United States presidential election.

[5] A myriad of different destabilizing scenarios could occur as a results of mass job automation. These scenarios would vary based on the speed of technological advancement, among other factors. Many such scenarios could be dramatically or even catastrophically destabilizing, including inciting military action or the nationalization of AGI development. Less dramatic, but potentially catastrophic possibilities, include the installment of a government opposed or agnostic to AI strategy/policy initiatives or the abandonment of previously ratified accords regarding AI safety.

Qualities of Super Wicked Problems	AGI strategy/policy	Narrow AI strategy/policy	
		Automation induced job loss	Militarization of AI
Time is running out: problems involving dynamic systems require the development of strategy/policy in a timely fashion, lest the problem become worse and its resolution more difficult.	A majority of experts expect the arrival of AGI within 100 years (Müller and Bostrom 2016), although there is a great degree of uncertainty regarding the timeline.	Currently, 47% of total US jobs are at a high risk of automation over the next two decades (Frey and Osborne 2017).	Autonomous drones are presently in use by the US military and billions are being allocated annually for the development of autonomous and semiautonomous weapons (Rosenberg 2016).
There exists no central authority: as a global issue the problem is broadly characterized as one of cooperation under anarchy. Moreover, decision makers within public institutions do not have complete control over the problem, thus resolutions require cooperation among not just states but across multiple economic sectors and political levels.	At present there exists no intergovernmental body dedicated to the pursuit of safe AGI, although a variety of organizations exist among the private and philanthropic sectors and in academia. The coordination among existing groups is poor.	There exists no intergovernmental organization specifically concerned with the effects of mass job automation. It seems that little economic incentives can be created to slow the development of automated systems.	There exists no intergovernmental body focused on the militarization of AI. Moreover, ongoing research increases the likelihood of an AI arms race and extends incentives to states furthering the militarization of AI. (NOTE: The topic is under discussion at the UN Convention on Conventional Weapons.)
Those seeking to end the problem are also causing it: affective stakeholders, either actively or passively, continue to contribute to the problem due to their inability to control all of the enabling factors.	AI technology is a rapidly growing sector, and a portion of all purchases can be expected to be used to further organizations' abilities to monetize such technology.	Economic incentives for the automation of jobs exist such that regulation would only hurt national economies. Only management of this problem is realistic.	Strategic incentives for the militarization of AI exist such that states are compelled to develop the technology to maturity prior to acceding to restrictions regarding use.
Hyperbolic discounting: decision makers and the public tend to make decisions that disregard strong evidence of negative impacts, rather, favoring short term benefits. Furthermore, such behavior persists despite overwhelming evidence of detrimental effects.	The development of precursors to AGI will substantially increase quality of life. It is unlikely that humans will choose to delay such improvements for the sake of longer term benefits. Such a delay would require unprecedented coordination.	Consumers and politicians make economic decisions based on short term benefits rather than considering long term consequences or preemptively preparing for or mitigating the consequent social impacts.	Policy makers opt for both short term and long term strategic benefits rather than weighing long term utility.

Table 2: the 4 central qualities of super wicked problems proposed by Levin et al. as applied to AI strategy/policy.

We also note that super wicked problems of class 2 are not limited to the two examples set forth here. Another problem entangled with AGI strategy/policy is the identification of a desirable outcome. Due to the inherent subjectivity of ascribing a single best future for the whole of humanity, this dimension of the problem is intractable. Solutions may exist such that the outcome need not be explicitly defined [Bostrom], but such solutions likely depend on the specifications of the system reaching AGI. This problem in and of itself arguably constitutes yet another independent super wicked problem entangled with AGI strategy/policy.

Levin *et al*. propose the use of a new epistemological approach, *applied forward reasoning*, when developing policy initiatives for the super wicked problem of climate change (Levin, Cashore et al. 2007, Levin, Cashore et al. 2012). In developing this technique, they draw heavily from literature on policy change and path dependency. Path dependency refers to the notion that previous conditions affect future conditions, or, in short, that in policy making 'history matters' (Howlett, Ramesh et al. 2003). They note the disentanglement of types of policies and the disentanglement of description from cause, from policy change and path dependency literature, respectively, as critical elements to addressing super wicked problems. Further, they draw on Page's review of path dependency literature and his proposed four pathways: lock-in, positive feedback, increasing returns and self-reinforcing (Page 2006). Their proposed approach contrasts that of the majority of path dependency scholars by looking forward, rather than backward, to illuminate the possibilities for how path dependencies can be employed constructively to foster desired policy outcomes in the futures.

Discussion

In this paper, we are not attempting to exhaustively assess the body of wicked problems literature in the context of AI strategy/policy. Rather, in this section, we simply would like to identify some implications on strategy and policy encountered when viewing the problem as a super wicked problem

Three formulations of criteria of wicked problems were presented. The formulation presented by Head was of particular interest in the context, and the simplest considering three key elements: complexity, uncertainty and divergence (or fragmentation). We first note that uncertainty is one of the most challenging components in efforts to develop AI strategy/policy. In this context it is of critical importance due precisely to one of the qualities of super wicked problems described by Levin *et al*.: time is running out. Thus, uncertainty about the timeline of the militarization of AI technology; about the ability to use AI to automate jobs; about the time to develop artificial general intelligence; poses a challenge not previously encountered in wicked or super wicked problems. Furthermore, the notion of divergence or fragmentation of ideas among stakeholders is inextricably tied to research on AI strategy/policy. Such fragmentation is inherent in human nature and history offers no examples of large scale global cooperation, however recent efforts toward climate change may offer the closest example.

Climate change is the only widely accepted example of a super wicked problem, and in many ways appropriate analogies can be made with AI strategy/policy problems (hopefully in ways that benefit society). There are, however, various ways in which the two problems diverge entirely. Perhaps the most significant of these conflicting dimensions involves the starkly different incentive structures of each. In the case of climate change, all stakeholders benefit from cooperation; in the case of AI strategy/policy such is not the case.

Of the organizational literature regarding wicked problems discussed here, a common theme was the necessity of organized and deliberate communication. Camillus proposed ideas to this effect and Conklin thoroughly details the mechanics and value of a technique embodying this notion. Moreover, Head *et al.* also suggest such strategies for dealing with wicked problems (Head and Alford 2015). Bostrom, in his recent preliminary analysis of openness in AI development, discussed virtues of specific forms of openness (Bostrom 2017). He concluded that openness regarding control methods and risk analysis was unambiguously good, given it be responsibly censored for other sensitive information, and that openness regarding values, goals and governance structures was generally welcome as well. Based on the work discussed in this essay, and particularly that of Camillus and Conklin, we concur entirely. Moreover, we suggest that regarding these dimensions of openness we not only develop policy initiatives, but actively work to develop a framework and digital platform for enabling meta-dialogue mapping available to all stakeholders[6]. We further suggest that this be maintained and supported by a collaborative effort such as the Partnership on AI or an intergovernmental agency[7]. Moreover, we suggest emphasizing the necessity of organized and deliberate communication within organizations regarding AI strategy/policy. In these scenarios, techniques such as dialogue mapping are likely to be of significant value[8]. Furthermore, it is important to remain cognizant of the effects of path dependency that can trickle up from implementation smaller policies.

One significant implication of super wicked problems on the AGI strategy/policy problem is to underscore the significance of path dependencies and to highlight their potential constructive roles in disentanglement. The dangers of path dependencies for these problems are well known, and congruent with the tenth property of wicked problems as described by Rittel & Webber: the planner has no right to be

[6] *i.e.* all of humanity.

[7] *e.g.* the intergovernmental panel on climate change (http://www.ipcc.ch/).

[8] This is to suggest, that, broadly, organizational management literature regarding wicked problems should be considered by involved organizations.

wrong. In this light, Flynn's suggestions for motivated individuals to not pursue AGI strategy/policy research could be dangerous, by inhibiting inclusion, and pushing undeterred individuals to pursue research in the domain without guidance or acceptable peer-review. This could lead to what Page describes as *lock-in*, which begins a path of policy interventions that are durable from the onset, and thus more difficult to alter. We suggest that those interested in AI strategy/policy research be supported insomuch that they are not excluded from related discourse which could lead to misinformed opinions being disseminated etc.

Finally, we note that Ackoff's classification of problem solving outcomes – *i.e.* problems as being either solved, resolved or dissolved – is worth considerating in the context of the unprecedented *mess* that is the problem of AI strategy/policy. Specifically, and due to the extraordinary wickedness of the problem, it is important to consider these forms of problems solving when developing AI strategy/policy for the myriad of constituent problems. In striving for the best possible solution, we must be thrifty with our resources and be cognizant of when resolving a problem rather than solving it is acceptable. This is to say that some of these problems are more tangled than others and thus some problems can affect the broader solution more than others – we must ensure that we do not use resources to solve problems when satisficing is acceptable

Conclusion

This essay explored the problem posed by AI strategy/policy research in the context of wicked problems. The broader problem was deconstructed into a hierarchy of multiple super wicked problems and a larger number of wicked problems, which, for the sake of brevity were ignored. The qualities of super wicked problems were explored with respect to the two hierarchical classes of super wicked problems. Based on this exploration, the problem of AGI strategy/policy was proposed as being more complex than other super wicked problems, itself being entangled with other super wicked problems of less significant impact potential. A discussion ensued, exploring the implications of identifying AI strategy/policy research as a super wicked problem and analyzing techniques for addressing wicked problems in the context of AI strategy/policy.

We wish to note that this paper is by no means intended to be an exhaustive exploration of the implications of wicked problems and super wicked problems on better understanding and addressing the gamut of research problems related to AI strategy/policy. Rather, it was intended as the opposite, merely an introduction to an alternate framework through which we may view this complex, entangled web of problems. Future work regarding AI strategy/policy as a super wicked problem will be critical to the disentanglement of research questions in AI strategy/policy.

In concluding remarks we wish to remind the reader of Rittel & Webber's final criterion in their list of wicked problem properties: the planner has no right to be wrong.

Acknowledgements

For discussion and comments we are grateful to David Paradice and Christy Manning.

References

Ackoff, R. L. (1974). "Redesigning the future." New York: 29.

Ackoff, R. L. (1978). The Art of Problem Solving Accompanied by Ackoff's Fables, John Wiley & Sons.

Bostrom, N. (2017). "Strategic Implications of Openness in AI Development." Global Policy **8**(2): 135-148.

Camillus, J. C. (2008). "Strategy as a wicked problem." Harvard business review **86**(5): 98.

Churchman, C. W. (1967). Guest editorial: Wicked problems, JSTOR.

Conklin, J. (2006). Dialogue mapping: Building shared understanding of wicked problems, Chichester, UK: Wiley Publishing.

Flynn, C. (2017). "Personal thoughts on careers in AI policy and strategy."

Frey, C. B., T. Berger and C. Chen (2017). "Political Machinery: Automation Anxiety and the 2016 US Presidential Election."

Frey, C. B. and M. A. Osborne (2017). "The future of employment: how susceptible are jobs to computerisation?" Technological Forecasting and Social Change **114**: 254-280.

Head, B. W. (2008). "Wicked problems in public policy." Public Policy **3**(2): 101.

Head, B. W. and J. Alford (2015). "Wicked problems: Implications for public policy and management." Administration & Society **47**(6): 711-739.

Howlett, M., M. Ramesh and A. Perl (2003). Studying public policy: Policy cycles and policy subsystems, Oxford University Press Oxford.

Lazarus, R. J. (2008). "Super wicked problems and climate change: Restraining the present to liberate the future." Cornell L. Rev. **94**: 1153.

Levin, K., B. Cashore, S. Bernstein and G. Auld (2007). Playing it forward: Path dependency, progressive incrementalism, and the 'Super Wicked'problem of global climate change. International Studies Association 48th Annual Convention. Chicago, February.

Levin, K., B. Cashore, S. Bernstein and G. Auld (2012). "Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change." Policy sciences **45**(2): 123-152.

Müller, V. C. and N. Bostrom (2016). Future progress in artificial intelligence: A survey of expert opinion. Fundamental issues of artificial intelligence, Springer**:** 553-570.

Page, S. E. (2006). "Path dependence." Quarterly Journal of Political Science **1**(1): 87-115.

Rittel, H. W. and M. M. Webber (1973). "Dilemmas in a general theory of planning." Policy sciences **4**(2): 155-169.

Rosenberg, M., Markoff, John (2016). At Heart of US Strategy, Weapons That Can Think. New York Times.