

Human Caused Bifurcations in a Hybrid Team—A Position Paper

Ira S. Moskowitz¹ and William Lawless²

¹Information Management and Decision Architectures Branch
Code 5580

Naval Research Laboratory, Washington, DC 20375
&

²Paine College
Department of Mathematics, Sciences and Technology
& Department of Social Sciences
Augusta, GA 30901

Abstract

We consider the effects of a human in the loop with respect to dramatic behavior in a team environment in this position paper. This dramatic behavior is captured mathematically as a jump in behavior. We cite recent examples and discuss earlier work in the cognitive sciences. We consider the problem in light of network science.

Introduction

In this position paper, we take the perhaps controversial view that previous work done in the cognitive and psychological sciences using catastrophe theory can be used to analyze and understand the behavior of hybrid teams (composed of non-humans and humans). We postulate, that the non-human elements of a team act in a continuous and gradual manner. However, the human elements of a team may behave in a singular manner, resulting in bifurcated behavior, which is the hallmark of catastrophe theory.

We start this paper with two recent examples of how a human member of a hybrid team caused a massive and almost instantaneous jump in team performance. Next, we discuss how various graph topologies can either amplify or diminish singular-type human behavior in a hybrid team. We next discuss results from the literature showing how visual cognition can be modeled as the simple “fold catastrophe” first proposed by René Thom (Thom 1975).

In accordance with the symposium guidelines, we present this paper in the hopes of generating comments and critiques from the community.

We consider the effects of a human in the loop with respect to dramatic behavior in a team environment. This dramatic behavior is captured mathematically as a *bifurcation* (Arnol’d 1972; Nicolis and Prigogine 1989) (or more specifically a *catastrophe* in certain situations). For us a hybrid

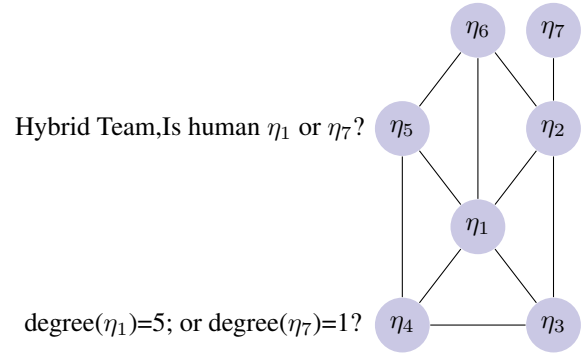


Figure 1: Various node degrees.

team is arbitrarily composed of humans, computers, machines, and robots. Our assumption, which is open to debate, is that the non-human elements of the team perform in a continuous, and possibly differentiable manner. The human however, due to his/her “aha” moment, and/or the vagaries of the human mind may, in certain circumstances, jump in its behavior space to a point that is not in its local neighborhood. This is the cause of the bifurcation.

Furthermore, how the team is interconnected as a small network or graph can either amplify the bifurcated behavior of the human or cause it to dampen. For example, if the human is a node in the graph with a very high node degree (Mieghem 2006), its odd behavior can greatly influence team performance. Whereas if the human is a node of low degree on the periphery of the graph, any jump in behavior may have only minimal impact upon team performance.

We realize that any mathematical modeling of a hybrid team is a hard problem, with many variables and unknown factors. We do not present this position paper as “the answer”, rather we present it as food for thought and discussion.

Furthermore, if the human in the team is outside a bottle-

neck, their catastrophic behavior may not affect the team regardless of node degree (Fig. 1). This can be well-analyzed using Cheeger’s constant (Donetti, Neri, and Muñoz 2006, Sec. 2).

The point is if the human, or in fact other team entities act in an irrational manner this affect can either be amplified or reduced by where the “bad player” is in the team. In fact the strength of this chaotic behavior can be measured probabilistically and/or by various tools from network science, such as the spectral radius, to gauge the affect upon team performance (Mieghem et al. 2011). We feel that this observation is worthy of further study and intend to research it.

Examples

Consider the recent Germanwings airlines crash (Smith 2015). The plane was functioning well, the software was functioning well, air traffic control was functioning well. All of these machines, humans, and software devices were part of the hybrid team which included the cockpit crew that was flying the plane. However, the co-pilot acted bizarrely and created a critical point in team performance. This critical point caused a massive jump in the hybrid team performance and the plane crashed. Note that the co-pilot in graph theoretical terms was a node of a very high degree. That is, the co-pilot’s actions had a great affect upon the team performance. The bifurcation came about due to the fact that the team performance was following a path in phase space of very stable behavior, and then it dramatically jumped to unstable behavior. Note that we use the term critical point to refer to where the function in question has less than maximal rank of its differential (Jacobian matrix) (Golubitsky, Keyfitz, and Schaeffer 1981; Seydel 1988). In fact, the differential may not even be defined at that point. It is a singularity of the Jacobian matrix not to be confused with singularities with respect to black holes (and infinite magnitude behavior). However, approaching a singularity may produce a large mechanical advantage for a inverted Jacobian that may need to be avoided (e.g. the Jacobian matrix relating joint velocities to end effector velocities).

Unfortunately, we have another recent example with respect to the Amtrak derailment (Nixon 2015). The influence of the human on the team is at the heart of the matter. What if the rest of the team instead of being a slave to the human, could in fact quickly override the human? That is, what if after a hybrid team arbitrarily composed of humans, robots or machines has been trained to perform perfectly as designed by its human masters, the machine knowledge used operationally to also measure the performance of the team including its human-supervisor-in-the-loop? This inverted application of measuring the performance of the human controller by including the human in the loop is at the heart of a symposium for the forthcoming AAAI conference at Stanford in the Spring 2016. Although controversial, examples of an application could have been used to override the co-pilot who flew the Germanwings aircraft into mountains, committing suicide while killing all aboard (Smith 2015); overriding the train engineer controlling the train who allowed his train to

run above the speed-limit of the rail tracks in the Northeastern corridor, killing many (Nixon 2015); and the possibility of a car being able to override its driver to prevent an accident (i.e., with a crash avoidance system).

Cognitive Background

Quoting from (Moskowitz, Longdon, and Chang 2002)

Catastrophe theory was developed in the 1970’s by ... René Thom (Thom 1975). In some sense, catastrophe theory was the unsuccessful precursor to chaotic dynamical systems. As the name implies, catastrophe theory models discontinuities in a system’s behavior, e.g., when does a dog decide to bark, what is the difference between genius and insanity, when does the bubble burst on Internet stocks, etc.? In short, it shows how discontinuities can describe certain aspects of continuous natural systems...

Note that Thom’s initial examples concentrated on a dog’s cognitive process—the bifurcation point where it goes from peaceful to attacking. van der Maas (van der Maas and Moleenaar 1992) has shown how catastrophe can be applied to the discontinuity hypothesis of Piaget (Piaget and Inhelder 1969). In this short position paper we concentrate on the work done on visual perception with respect to catastrophe theory (Poston and Stewart 1978; Stewart and Peregoy 1983) related to the earlier work in perception originating with Fisher’s famous illustration (Fisher 1967; Fisher 1968; Attneave 1971).



Figure 2: Original Image: Man’s face or Sitting Woman?

In (Poston and Stewart 1978; Stewart and Peregoy 1983) Fisher’s work was parametrized as a sequence of images that go from clearly a man’s face to clearly a sitting woman (Figs. 2,3). This is similar to the human behavior that we wish to capture in our hybrid team analysis (e.g., shifting from a normal co-pilot to a suicidal co-pilot). The parametrized images are mathematically modeled by the energy function V (Stewart and Peregoy 1983)

$$V(x) = x^3 + ax. \quad (1)$$

The first and second derivatives are

$$V'(x) = 3x^2 + a \quad (2)$$

$$V''(x) = 6x \quad (3)$$



Figure 3: Behavior of the parameter a : NEGATIVE $\rightarrow 0 \rightarrow$ POSITIVE
 Parametrized Sequence of Images: left is $a < 0$, (original image) “middle” is $a = 0$, right is $a > 0$

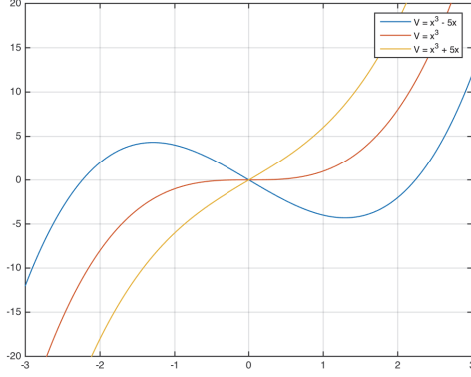


Figure 4: Energy Functions

Therefore it is easy to see that $V(0) = 0$, $V(x)$ is concave down (up) for $a < 0$ ($a > 0$), and that $V(x)$ has critical points ($V'(x) = 0$) only when $a < 0$. Let us look at the critical points in detail. The critical points are

$$x = \pm \sqrt{-a/3},$$

$$\text{with } V(\pm \sqrt{-a/3}) = \mp 2 \left(\frac{a}{3} \right)^{3/2}.$$

At $a = 0$ we have the simple cubic x^3 , and as $a > 0$ goes to infinity, $V(x)$ becomes more linear near zero.

For the plot of $V(x) = x^3 - 5x$ shown in blue in Figure 3 we see the two critical points and their corresponding V values. The positive critical point 1.29 creates a local minimum -4.3, and the negative critical point -1.29 creates a local maximum 4.3.

Note (for $a < 0$) that $(\sqrt{-a/3}, -2(\frac{a}{3})^{3/2})$ corresponds to a local minimum energy state. However, if we look at the closed system as having fixed energy the local minimum is an equilibrium point with maximal entropy. This corresponds with the theory of *entropic force* F where the “right” response is in the direction of maximal entropy increase (along the gradient of the entropy S in state X , where T is the system temperature) (Wissner-Gross and Freer 2013).

$$F = T \nabla_X S(X) \quad (4)$$

Consider what happens to the local minimum $(\sqrt{-a/3}, -2(\frac{a}{3})^{3/2})$ as a increases. The deep well gets more and more shallow until, at $a = 0$ it ceases to exist. For a very negative, the choice of the image is man, as a increases there is some possibility that one could interpret the image as a woman. Finally at $a = 0$ we reach

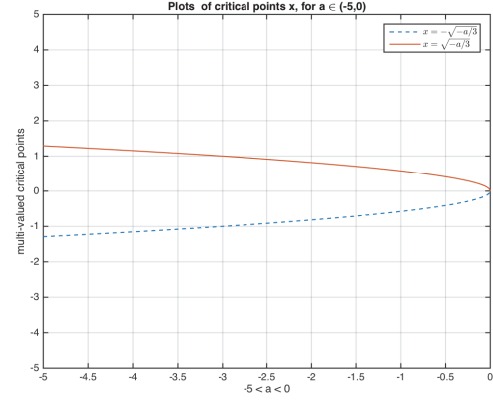


Figure 5: Stable (bottom) and Unstable (top) Critical Points

a catastrophe and there is only one choice, woman (Stewart and Peregoy 1983).

The behavior exhibited by this system is called the *fold catastrophe*. This name comes about from the illustrated behavior.

When $a = 0$ we have a catastrophe, we go from two critical points to none, see Figure 4. This is a phase transition and a jump bifurcation. This jump or bifurcation is the jump caused by a human. We plan to expand upon this idea and see how human bifurcated behavior affects a team as a whole in future work.

Conclusion

We take the position that using catastrophe modeling of human behavior can be used to see how hybrid teams perform. At this point we have no experiments to back up our assertion. However, we feel it is an area worth discussing with the participants at this symposium (in fact such speculative thinking is part of the reason for the being of this symposium).

We conclude this paper by paraphrasing from (Stewart and Peregoy 1983):

Since the use of catastrophe theory in modeling psychological phenomena was introduced by Zeeman in 1971, it has been met variously with enthusiasm (...), cautious experimentation (...), and outright dismay (...).

We expect the same for our ideas.

We are cognizant that our ideas are just that: ideas. We plan in future work, if this position paper is well-received, to expand upon what we have proposed here. Of course, the feedback from the symposium attendees will allow us to hone our ideas into a better product.

References

- Arnol'd, V. I. 1972. Lectures on bifurcations in versal families. *Russian Mathematical Surveys* 27(5):54.
- Attneave, F. 1971. Multistability in perception. *Scientific American* 225(6):62–71.

- Donetti, L.; Neri, F.; and Muñoz, M. A. 2006. Optimal network topologies: expanders, cages, Ramanujan graphs, entangled networks and all that. *Journal of Statistical Mechanics: Theory and Experiment* P08007.
- Fisher, G. 1967. Preparation of ambiguous stimulus materials. *Perc. Psychopos.* 2:421–422.
- Fisher, G. 1968. Ambiguity of form: Old and new. *Perc. Psychopos.* 4:189–192.
- Golubitsky, M.; Keyfitz, B. L.; and Schaeffer, D. G. 1981. A singularity theory analysis of a thermal-chaingbranching model for the explosion peninsula. *Communicaitons on Pure and Applied Mathematics* 34:433–463.
- Mieghem, P. V.; Stevanoic, D.; Kuipers, F.; Li, C.; and van de Bovenkamp, R. 2011. Decreasing the spectral radius of a graph by link removals. *Physical Review E* 84:1–12.
- Mieghem, P. V. 2006. *Performance Analysis of Communications Networks and Systems*. Cambridge University Press.
- Moskowitz, I. S.; Longdon, G. E.; and Chang, L. 2002. *A New Paradigm Hidden in Steganography*. CRC Press. chapter 27 of *The Privacy Papers*, ed. R. Herold, 331–349. (Also appeared in Proc. NSPW 2000).
- Nicolis, G., and Prigogine, I. 1989. *Exploring Complexity*. NY: Freeman.
- Nixon, R. 2015. “No evidence that amtrak engineer used phone during crash, safety board says”, NY Times.
- Piaget, J., and Inhelder, B. 1969. *The Psychology of the Child*. NY: Basic Books.
- Poston, T., and Stewart, I. 1978. Nonlinear modeling of multistable perception. *Behavioral Science* 23:318–334.
- Seydel, R. 1988. *From Equilibrium to Chaos*. NY: Elsevier.
- Smith, P. 2015. “Why pilots still matter”, NY Times.
- Stewart, I., and Peregoy, P. 1983. Catastrophe theory modeling in psychology. *Psychological Bulletin* 94(2):336–362.
- Thom, R. 1975. *Structural Stability and Morphogenesis*. *Fr. ed* 1972. Reading, MA: W.A. Benjamin.
- van der Maas, H. L., and Molenaar, P. C. 1992. Stagewise cognitive development: An application of catastrophe theory. *Psychological Review* 99(3):395–417.
- Wissner-Gross, A. D., and Freer, C. E. 2013. Causal entropic forces. *Physical Review Letters* 110:168702.