

Black Swans, Gray Cygnets and Other Rare Birds

Edwina L. Rissland

Department of Computer Science
140 Governors Drive
University of Massachusetts
Amherst, MA 01003-9264
rissland@cs.umass.edu

Abstract. Surprising, exceptional cases — so-called black swans — can provoke extraordinary change in the way we do things or conceptualize the world. While it is not unreasonable to be surprised by a black swan, to be surprised by subsequent cases that are similar enough that they might cause the same sort of upheaval is unforgivable. The problem is how to reason about these almost novel, not totally unforeseen, subsequent cases that I call **gray cygnets**.

1 Introduction

From time to time there are unanticipated, exceptional cases that provoke extraordinary change in the way we do things or conceptualize the world. It is popular today to call such events *black swans* (e.g., [24]). They can be either positive or negative in the sense that they can expand or contract a conceptualization or approach. Not only do black swans trigger great changes, but also they are usually outside our conceptualization of the world and thus, come to us unheralded and seemingly out of the blue. They are said to exist in the “fat tail” of concepts.

While it is not unreasonable to be surprised by a black swan, to be surprised by subsequent cases that are similar to the black swan — in fact, similar enough that they might cause the same sort of upheaval — is unforgivable. As is often said: fool me once, shame on you; fool me twice, shame on me. If an entity — person or machine — cannot learn not to be fooled by essentially the same circumstance a second, or a third time, it is hardly intelligent.

The problem is how to spot and reason about cases that are similar enough to the original black swan that their effect might be similarly significant and exceptional. These almost novel, not totally unforeseen, subsequent cases are what I call **gray cygnets**.

2 The Gray Cygnet Problem: Four Research Issues

There are several aspects of the gray cygnet problem that are particularly relevant to case-based reasoning (CBR) and for which CBR is particularly well-suited. These include:

1. **Similarity assessment.** How similar to a black swan is similar enough for an event to be considered a gray cygnet? This is the classic CBR concern of “similarity assessment”. Further, given a new problem situation, how should a highly similar past black swan case be weighed against a highly similar past run-of-the-mill case? Which case, if any, should guide the system’s response? This is the classic “best case selection” issue.
2. **Hypotheticals.** How can hypotheticals be used to anticipate gray cygnets and explore the space of possible future cases and/or boundaries of current models, assumptions, policies, conceptualizations, etc. For instance, hypos based on a black swan can be created by simplifying it, pushing it to extremes, softening it to be a near hit or near miss, etc. and these can be used to plumb the instance space, particularly the fat tail where real data is sparse.
3. **Re-representation.** *How — and when — should the representation of a concept, plan, model, policy, etc. be changed in response to a black swan and gray cygnets?* For instance, how often should an underlying causal model be revised? What are the trade-offs between aggressive re-modeling with its risks of unstable model oscillations versus a “lazy” approach with its risks of using a model no longer valid? Should notions of what is typical or prototypical be revised? Should the validity of past analyses and solutions be revisited or even called into question?
4. **Analogical Mapping and Explanation.** *How should lessons learned from the black swan be mapped over or applied to a gray cygnet?* For instance, should one reason analogically. Should a new — “other” or anomaly — class that includes cygnet and swan be created, perhaps with minimal deep reasoning. What is the role of deeper explanation, and how should the exceptional treatment of a black swan and/or gray cygnets be explained?

Understanding the gray cygnet problem is important for several reasons. First, it is necessary for building intelligent systems that are not only robust in ever changing and challenging contexts but that can also be pro-active in such domains. Given that change is ubiquitous, it is necessary to understand how systems, including CBR systems, can learn in response to surprising, exceptional, and provocative events. Second, understanding the gray cygnet problem can shed insights on what is meant by concepts and how to represent them. Most concepts in the world are “messy” in the sense that they cannot be represented by universally valid necessary and sufficient conditions — compared with mathematical concepts, say — but yet there is often a great deal of regularity in them. This is the view pioneered by Rosch and colleagues in psychology research on notions of “typicality”, “prototypes”, and “atypical” instances (e.g., [22] , [21]). How do swans and cygnets participate in such conceptualizations?

Similar viewpoints are found in legal philosophy concerning the differences between reasoning, and the nature of cases themselves, in the so-called “core of settled meaning”, where there are regularities and general heuristic truths and in the “penumbra”, where the cases are *sui generis* in the sense that each is somewhat unique and does not fit in the mold of cases in the settled core [5]. In CBR, an

analogous concern is the use of hybrid representations and reasoning, like rules or models and cases [20],[2],[7],[13], and more generally, mixed model-plus-extensional approaches for representation, such as frames or logic plus actual exemplars, which has long been of interest to me [18].

3 Examples of Black Swans and Gray Cygnets

Appellate law offers many examples of the scenario in which a black swan case creates an initial exception under which a line of gray cygnets emerge, so many so that eventually the accepted rule or approach becomes obsolete. For instance, the contract case *Thomas and Wife v. Winchester* (discussed in [12]) was the black swan that opened the way for a whole bevy of gray cygnets that eventually caused reversal of the “privity of contract” doctrine (that required a direct relationship between the manufacturer of a product and a party harmed by the product) and heralded legal changes now reflected in modern consumer rights law (that allows a consumer to recover damages for harm caused by a product, such as a car, bought through an intermediary like a dealer or store).

Another legal example is the 1925 prohibition era case of *Carroll v. United States*, 267 U.S. 132 (1925). This black swan created the so-called “automobile exception” under which police do not need a warrant to search a car. This exception flew in the face of the longstanding existing Fourth Amendment jurisprudence that said a warrant was necessary. The *Carroll* case created an exception that grew large enough to drive a truck through, and created a bifurcation in the concept of constitutionally valid searches — essentially the dichotomy of homes versus cars — that persisted for nearly 50 years, when the (not unimaginable) case of a house on wheels (i.e., a truck with a camper compartment) presented itself [17],[16]. Imagine if a hypo of a car used in a homelike way — say, for camping — had been raised in 1925. This is not an unimaginable hypo: think of Conestoga wagons, an example of a vehicle used for both transportation and homelike activities. Would the automobile exception have occurred or been justified as it was?

Black swans can occur in many disciplines, for instance, crashes or burst bubbles in economic markets. Financial black swans are the main focus of the Taleb book. The Great Depression is often treated as a black swan since it created great upheaval and was deemed a rare event. It is yet to be seen whether the recent “crash” in October 2008 will earn the title; it certainly created upheaval and change, but whether it was a true surprise is an open question.

Black swans can occur in medicine, often with quite negative results. For instance, the so-called Libby Zion case precipitated a drastic change in the way patients are treated in the Emergency Room and residents are trained and practice medicine. Libby Zion, an 18 year old college student, was seen in the ER for fever and mysterious jerking movements, admitted and sedated with an opiate (meperidine), and subsequently physically restrained and given further sedation (haloperidol) when she became agitated again. The next morning she died from cardiac arrest triggered by an alarming (107°) fever. While the cause of death was not determined definitively, it is suspected to have been

from drug interactions with the anti-depressant (phenelzine) she was taking. The only doctors that saw her were residents in training and they were routinely exhausted from long shifts. Training and protocols are now markedly different. ER physicians pale at the thought of creating a Libby Zion cygnet.

Even mathematics and science are not immune to black swans. Lakatos [10] discusses various treatments of some surprising counter-examples (e.g., Kepler's stellated dodecahedron) in the history of mathematics concerning Euler's formula. Weintraub [25] discusses examples from the long history of what heavenly objects are considered planets. For example, in the days of the early Greeks, Earth was not classified as a planet but the Sun was; the Copernican revolution switched their statuses. Pluto, discovered in 1930, was originally classified as a planet. Since then however, the discovery of a plethora of Trans-Neptune Objects has played havoc with the planet concept. Many astronomers feel that there are now far too many Pluto-like objects for them all to be called "planets"; so, to handle them, a new category called Plutinos was invented. If consistency is to be maintained, Plutinos along with Pluto should all be in or out of the concept class of planet. So the choice is either Pluto flips to being a non-planet or there are scores of new planets. Many feel that this enlarges the concept class too far.

In these examples, a drastic change is first heralded by a surprising, exceptional case (the black swan) that is then followed by a whole flotilla of subsequent, similar, exceptional cases (the gray cygnets). If one thinks of the set of positive instances of a concept, model, or a rule application as a disc, the black swan creates a puncture in it and the gray cygnets enlarge the hole. Eventually, there might not be much left of it. There might even be exceptions to the exceptions: positive islands within negative holes. Black swans can also occur near the boundary region and cause the concept to stretch or contract, thus flipping the classification of some nearby, borderline instances.

Gray cygnets, while perhaps not as "black" as the original event or case in the sense that they are obviously no longer as novel, can still be exceptional and have far-reaching ramifications, for instance by establishing a line of cases that solidify concept change. In precedent-based domains like law, gray cygnets tend to cite back to the original black swan case as justification, and once an exception has been allowed (e.g., not dismissed as a wrongfully-decided mistake), it is inevitable that more cases will be treated similarly. Eventually, this flock of exceptions might become the norm, when, as is often said, 'the exception swallows the rule'. In planning tasks, cygnets can indicate problems in underlying assumptions (e.g., unanticipated co-occurrences of circumstances).

4 Recognition of Swans and Vigilant Monitoring for Cygnets

There are two separate problems involving swans and cygnets: (1) recognizing a black swan, and (2) vigilant monitoring for subsequent, similar gray cygnets.

With regard to recognition, I note that it is often difficult to recognize a black swan event contemporaneously with its occurrence, especially in weak theory

domains, the forte of CBR, where one might need to await future commentary since it is often only in retrospect that there is recognition that one has occurred. This might have been the scenario in Levi example; the *Winchester* case was probably not considered a black swan when it was decided. It is often only when subsequent cases cite back to the black swan case as the enabling precedent — often with the rubric “landmark” case or use of its name as a shorthand to stand for a new way of thinking — that a black swan is truly recognized as such. In other domains, like planning, recognition might be more timely because the suspected black swan causes an immediately obvious surprise like an unanticipated failure or success [8]. That a case badly fits current expectations can lead to the hypothesis that a black swan event has occurred or is occurring [9]. However, it is typically only through closer analysis that one understands why it is exceptional, and this often requires more knowledge than just the case itself.

Once a surprising potentially paradigm-threatening case has occurred — even if not dubbed a black swan — one should not be snookered by similar, potentially disruptive, subsequent cases, that is, gray cygnets. One needs to be on the alert for them and invest in the reasoning required to understand their possible ramifications and monitor the case-stream for similar occurrences. While in some tasks, there is a rich enough domain knowledge to support pre-emptive measures, in many there is not, and watchful vigilance is the only alternative. That some cases turned out to be much ado about nothing and that vigilance wasted some resources does not, for me, obviate this approach since the cost-benefit trade-offs in many domains place a high penalty on missing cygnets. The worst possible approach in my opinion would be to dismiss the original black swan or subsequent cygnets as mere outliers or mistakes. This would open one up to making the same mistake twice, the hallmark of the opposite of an intelligent system.

5 Hypotheticals and Synthetic Cygnets

Creating hypotheticals is a way to approach the gray cygnet problem proactively. Hypos help perform a knowledge-based search of the case space (cf, statistical sampling) and provide useful ersatz data when real data is sparse. They can be used to probe the current conceptual regime. A hypothetical gray cygnet can be particularly telling.

One way to create synthetic cygnets is to take a black swan as a starting “seed” case and modify it heuristically in ways that are important in the domain to create derivative hypos or an entire “constellation” of them. Methods based on “dimensions” have been used to implement this approach in HYPO-style systems as well as analyze how experts do it [19],[17]:

- (i) Make a case weaker or stronger along a dimension
- (ii) Make a case extreme
- (iii) Enable a near miss (near negative)
- (iv) Dis-able a near win (near positive)

- (v) Add a closely coupled aspect
- (vi) Add potentially conflicting aspects
- (vii) Weaken a clear win (solid positive) to be a near win.

For instance, in the warrantless search area, major dimensions are “expectation of privacy” and “exigency”. One can strengthen a fact situation on the first dimension by making a vehicle more homelike by adding features that indicate use of the vehicle’s interior for living (e.g., privacy curtains, bedding, cooking equipment). One can weaken the exigency dimension by compromising a vehicle’s mobility (e.g., by having it hooked up to utilities in an RV park, parked up on blocks). One can create a conflict hypo by making a situation extremely strong on both, for instance, a fully equipped RV capable of traveling at highway speeds. Such hypos were actually used in actual Supreme Court oral arguments by the Justices to explore the ramifications of potential analyses and decisions [17].

Besides dimension-based techniques, there are others used in the CBR community, for instance those based on distance metrics. For instance, using the k-NN metric, one can create a mid-point hypo or the “nearest unlike neighbor” or “NUN” [15],[14].

6 Responsive Re-representation

Once a black swan and gray cygnets have been discovered, the issue is how to revise the conceptual regime. Some responses can mimic those discussed by Lakatos, for instance, “surrender” (reject the old regime or model entirely) or “monster barring” (bar them from being a legitimate case in the instance space).

In Anglo-American common law, the standard response is to add the black swan to the case base and then simply allow exceptional gray cygnet cases to accrete in the case base so long as they are justified in a precedent-based manner (e.g., citing the black swan). The law does not explicitly remove from the case base cases that are discredited by the black swan, or that rely on them. They are still in the case base, but it is up to legal practitioners to know that they might no longer be good law and of little precedential value. Eventually, the set of black swan plus gray cygnets can dominate the case-base or be declared by some (appellate) court to be the correct approach. Levi’s privity of contract example and the automobile exception examples illustrate these phenomena.

In statutory law where there are rules as well as cases, the black swan and its cygnets can create exceptions, temporary updates, or implicit work-arounds to the statutory rule. Eventually the body that wrote that statute (e.g., Congress) can re-write the rule to reflect the swan and cygnets in a revised statute. In some cases, like tax law, new regulations and advisories can be issued to say how such cases are to be handled. My lab’s hybrid rule and case-based system CABARET used cases in this way (see [20], [23]).

In a domain with an explicit theory, an appropriate theory revision mechanism can be used. For instance, if the domain theory is an ontology, it can be revised by splitting or removing subclasses. If there is a causal model, causal pathways

can be added or removed, parameter ranges restricted or expanded, new variables introduced, etc. Of course, such revisions can be difficult because they can involve age-old AI problems like credit assignment.

In my lab, we continue to investigate *hybrid CBR and representations for concepts, domain models, policies, etc.* We use (a) a general summary representation PLUS (b) a set of exemplars, called the **E-set**, a selective subset of the case knowledge base. The summary representation might take the form of definitional rules, HYPO-style dimensions, prototypes, or causal models. The E-set includes both positive and negative instances and is the repository for “atypical” or “penumbral” or “surprising” cases. Black swans and gray cygnets can be captured and stored as members of the E-set. Re-representation can occur through accretion of cases in the E-set and/or explicit revision of the summary aspect. Management of the E-set is an interesting problem in itself. For instance, when a summary representation is revised, should the old E-set be retained for future use: totally, in part, not at all?

Finally, mapping over the lessons learned from a black swan or gray cygnet involves analogy and explanation. In addition to HYPO-style analogical reasoning [1], there is much relevant work by Gentner and Forbus on “structure mapping” that can provide an approach to analogy [4],[3]. It was fundamental to Branting’s GREBE project in a statutory legal domain [2]. A different and also highly effective approach was used in Hammond’s CHEF [6]. Of course, there is much powerful research by Roger Schank, David Leake and others on explanation involving cases [11].

References

1. Ashley, K.D.: Modeling Legal Argument: Reasoning with Cases and Hypotheticals. MIT Press, Cambridge (1990)
2. Branting, L.K.: Building Explanations from Rules and Structured Cases. *International Journal of Man-Machine Studies (IJMMS)* 34(6), 797–837 (1991)
3. Falkenhaimer, B., Forbus, K.D., Gentner, D.: The Structure-Mapping Engine: Algorithm and Examples. *Artificial Intelligence* 41(1), 1–63 (1989)
4. Gentner, D.: Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science* 7(2), 155–170 (1983)
5. Hart, H.L.A.: Positivism and the Separation of Law and Morals. *Harvard Law Review* 71, 593–629 (1958); reprinted in Hart, *Essays in Jurisprudence and Philosophy*. Oxford (1983)
6. Hammond, K.J.: Case-Based Planning: Viewing Planning as a Memory Task. Academic Press, London (1989)
7. Hastings, J., Branting, L.K., Lockwood, J.: CARMA: A Case-Based Rangeland Management Advisor. *AI Magazine* 23(2), 49–62 (2002)
8. Horling, B., Benyo, B., Lesser, V.: Using Self-Diagnosis to Adapt Organizational Structures. In: *Proc. AGENTS 2001, Montreal* (2001)
9. Horvitz, E., Apacible, J., Sarin, R., Liao, L.: Prediction, Expectation, and Surprise: Methods, Designs, and Study of a Deployed Traffic Forecasting System. In: *Proc. UAI* (2005)
10. Lakatos, I.: *Proofs and Refutations*. Cambridge (1976)

11. Leake, D.B.: *Evaluating Explanations: A Content Theory*. Lawrence Erlbaum, Mahwah (1992)
12. Levi, E.H.: *An Introduction to Legal Reasoning*. Chicago (1949)
13. Marling, C., Rissland, E.L., Aamodt, A.: Integrations with case-based reasoning. *Knowledge Engineering Review* 20(3), 241–246 (2005)
14. Massie, S., Wiratunga, N., Craw, S., Donati, A., Vicari, E.: From anomaly reports to cases. In: Weber, R.O., Richter, M.M. (eds.) *ICCBR 2007*. LNCS, vol. 4626, pp. 359–373. Springer, Heidelberg (2007)
15. Massie, S., Craw, S., Wiratunga, N.: Complexity-Guided Case Discovery for Case Based Reasoning. In: *Proc. AAAI 2005*, pp. 216–221 (2005)
16. Rissland, E.L.: AI and Similarity. *IEEE Information Systems* 21(3), 39–49 (2006)
17. Rissland, E.L.: Dimension-based Analysis of Hypotheticals from Supreme Court Oral Argument. In: *Proc. Second International Conf. on AI and Law (ICAIL 1989)*, Vancouver, BC, pp. 111–120 (1989)
18. Rissland, E.L.: *Epistemology, Representation, Understanding and Interactive Exploration of Mathematical Theories*. Ph.D. dissertation, Department of Mathematics, MIT (1977)
19. Rissland, E.L., Ashley, K.D.: Hypotheticals as Heuristic Device. In: *Proceedings Fifth International Conference on Artificial Intelligence (AAAI 1986)*, pp. 289–297 (1986)
20. Rissland, E.L., Skalak, D.B.: CABARET: Statutory Interpretation in a Hybrid Architecture. *International Journal of Man-Machine Studies (IJMMS)* 34, 839–887 (1991)
21. Rosch, E., et al.: Basic Objects in Natural Categories. *Cognitive Psychology* 8(3), 382–439 (1976)
22. Rosch, E., Mervais, C.B.: Family Resemblance: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7, 573–605 (1975)
23. Skalak, D.B., Rissland, E.L.: Arguments and Cases: An Inevitable Intertwining. *Artificial Intelligence and Law* 1(1), 3–44 (1992)
24. Taleb, N.N.: *The Black Swan*. Random House (2007)
25. Weintraub, D.: *Is Pluto a Planet?* Princeton (2007)