Report Number: WUCS-89-07

1989-03-08

Analogical Reasoning, Defeasible Reasoning, and the Reference Class

R. P. Loui

This paper attempts four things. It demonstrates the possibility of accounting for Russell-style and Clark-style analogical reasoning in an existing framework for statistical reasoning. It critically reviews the proposals made by Clark for defeasible analogical reasoning and shows how they can be understood better simply as defeasible reasoning. It argues that generalization from the single case is not as desirable as projection from the single case; the difference has to do with the defeasibility control strategy for statistical reasoning limited to a small number of cases.

[Department of Computer Science & Engineering](#) - Washington University in St. Louis
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

ANALOGICAL REASONING, DEFEASIBLE REASONING, AND THE REFERENCE CLASS

Ronald P. Loui

WUCS-89-07

March 1989

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

.

Analogical Reasoning, Defeasible Reasoning, and the Reference Class

R. P. Loui *
Dept. of Computer Science and
Dept. of Philosophy
Washington University
St. Louis, MO 63130
loui@ai.wustl.edu

March 8, 1989

Abstract

This paper attempts four things. It demonstrates the possibility of accounting for Russell-style and Clark-style analogical reasoning in an existing framework for statistical reasoning. It critically reviews the proposals made by Clark for defeasible analogical reasoning and shows how they can be understood better simply as defeasible reasoning. It argues that generalization from the single case is not as desirable as projection from the single case; the difference has to do with the defeasibility of the inference. Finally, it muses about the prospects for an appropriate control strategy for statistical reasoning limited to a small number of cases.

1 Introduction.

1.1 The Logical Problem of Analogy.

Analogy has been studied in a variety of activities by AI authors, notably in problem solving, in learning, and in common-sense reasoning (consider, for instance, [Winston80], [Kedar-Cabelli86], and [Gentner83]). An excellent recent review of this work can be found in Stuart Russell's thesis [Russell87].

---

*I have had useful discussions with Stuart Russell, Ben Grosof, Guillermo Simari, Josh Tenenberg, Henry Kyburg, John Pollock, Mike Wellman, Fahiem Bacchus, and Dana Nau on this subject in that temporal order. Thanks to Michael Anderson for leaving Stuart's thesis on the coffee table.

Most of the hopes to use analogy are as ampliative, unsound inference, but inference we are nevertheless willing to perform because of epistemological constraints – we don't know enough to do full induction. In the present treatment of analogy, I am equally concerned with situations in which our willingness to perform analogical reasoning derives from computational constraints. We may not want to include in our inductive reasoning as many cases as we are capable of recalling and considering. In the extreme case, we may even want to reason from a single case. It is appropriate to consider such reasoning from a single case to be analogy.

Among the many problems that seem to involve analogy, Russell distinguished the logical problem of analogy. This is the aspect of analogy of most interest here. How is analogical reasoning justified? Presumably, it is some form of induction, in the presence of added assumptions. A property is transferred from source case to target case in virtue of some shared properties. Apparently, there is inductive support for the possession of that property given the similarity. The interesting challenge is to render those assumptions in some underlying reasoning framework in such a way that the single source case contributes in the inference: it cannot be just that the probability of the transferred property given the shared properties is known to be high, because that makes the source case redundant.

Russell's approach was to justify analogical inference in a deductive framework. My approach will be to cast analogy as a special case of a logically sophisticated kind of statistical reasoning. Russell had examined as candidates for the explication of analogy less expressive frameworks for induction than the one considered here. Part of this exercise is to relieve his pessimism about the prospects of an inductive explication of analogy. We should welcome this relief because we all continue to think of analogy as inductive. Analogical reasoning so closely resembles inductive reasoning that it is a travesty to suggest that there can be no meaningful reduction. This paper hopes to correct any misapprehensions about the relation between inductive (or statistical) reasoning and analogical reasoning.

More importantly, basing analogy on statistical inference allows the capture of a wider class of common-sense reasoning. The study of reasoning from a small number of cases is the natural extension of analogy with a single case. Philosophers of induction may not see the point of reasoning from a small number of cases, just as they have often turned their noses to the special case of analogy. They might think just to run the induction mill on the cases available, whether they be one, three, or a thousand. But here's the point: we are sometimes unwilling to complete this computation. Our control strategy for arriving at interesting partial computations must be informed by what we know about the legitimacy of inferences from a small number of cases.

A different example of common-sense reasoning that extends analogical reasoning is studied by Peter Clark [Clark88]. Here, the problem is not with the multiplicity of cases, but with the multiplicity of shared properties. Clark's infer-

ences are interesting for a variety of reasons, one of which is that it uses analogy to amplify defeasible reasoning. It is Clark's ambitions that most clearly point out the need to base analogy in an appropriate statistical framework. What seems to be a complex use of analogy turns out to be simply the use of specificity in defeasible statistical reasoning; what seems to be an unquestionably sound maxim for transferring from source to target turns out to be improper in the most disturbing and unexpected statistical way.

1.2 Russell and Clark on Analogical Inference.

Russell's deductive account [Russell87] of analogical reasoning is at once authoritative and demonstrating exemplary clarity. Peter Clark's recent attempt to generalize the analogical reasoning paradigm to arbitrate among conflicting defeasible arguments is refreshing and exciting [Clark88]. These two papers represent two of the most formal accounts of analogical reasoning.

Russell is interested in inferences such as

Nationality DETERMINES *Language*
Language(Louis, French)

Nationality(Louis, France)

$\underline{Nationality(Antoinette, France)}$
thus,
Language(Antoinette, French)

by the analogy of the target, Antoinette, to the source, Louis. Louis's language, namely French, determines Antoinette's language to be French, in virtue of their important similarity. Russell justified this reasoning by taking the determination relation to mean, roughly,

$(w)(y)(z).$
$(Nationality(w,y) \land Language(w,z)) \rightarrow$
$\quad (x).Nationality(x,y) \rightarrow Language(x,z).$

This makes the inference valid in a first order system.

Peter Clark is interested in inferences such as

Over-Block(x) INDICATES $\neg Sand\text{-}At(x)$
Sand-Nearby(x) INDICATES *Sand-At(x)*

3

$Late\text{-}Fault(x)$ INDICATES $Sand\text{-}At(x)$
$Unfavorable\text{-}Environment(x)$ INDICATES $\neg Sand\text{-}At(x)$

$Over\text{-}Block(well_1)$
$Sand\text{-}Nearby(well_1)$
$Late\text{-}Fault(well_1)$
$Unfavorable\text{-}Environment(well_1)$
$Sand\text{-}At(well_1)$

$Over\text{-}Block(well_2)$
$Sand\text{-}Nearby(well_2)$
$Late\text{-}Fault(well_2)$
$Unfavorable\text{-}Environment(well_2)$
$\overline{\phantom{Unfavorable\text{-}Environment(well_2)}}$
thus,
$Sand-At(well_2)$

by the analogy of the target, $well_2$, to the source, $well_1$.

The resolution of the conflicting arguments in the case of $well_1$ determines the resolution of the conflicting arguments in the case of $well_2$. I will say that *Sand-Nearby* and *Late-Fault* are factors that indicate the ultimate conclusion. *Over-Block* and *Unfavorable-Environment* are factors that counter-indicate the ultimate conclusion. Note that in Russell's notation, we might have said that $Over\text{-}Block(x, true)$ determines $Sand\text{-}At(x, false)$, if $Over\text{-}Block$ had been the lone indicating factor. The easiest way to translate between Russell and Clark is to consider the case wherein Russell's relations are bivalent in the second argument (and determinations apply only when values are true).

Clark is interested in an additional behavior that I will call "monotonicity of conflict resolution." He thinks that the above inference should be supported even if

$Over\text{-}Block(well_2)$

were omitted as a premise about the target, and even if

$Sand\text{-}Nearby(well_1)$

were omitted as a premise about the source. If the target case contains more indicators of the conclusion and fewer counter-indicators than the source case, thinks Clark, then the conclusion should be preserved. Clark formalized this reasoning, but did not attempt to justify it, deductively or inductively.

4

1.3 The Reference Class

The account of statistical reasoning I have in mind is due to Kyburg [Kyburg61,74,82], though there may be similar reconstructions with Pollock's competing theory [83,84]. What is crucial about these theories is that they provide a logic for determining what is called "the reference class." The account of analogy I am considering could not be given in theories of enumerative induction, such as considered by Russell. Nor could it be given in a Bayesian theory of statistical inference in which determination of the reference class plays no role.

The problem of determining the reference class begins with explicit acknowledgement of multiple statistical sources. Essentially, choosing the reference class amounts to choosing the right statistical source on which to base inductive judgement. If we think of Neyman-Pearson statistics, the closest kin to the problem of choosing the reference class is the problem of testing for significant difference in two populations. But the statistician's logic is incomplete; most often, a single, homogeneous sampling population is assumed.

When there is information from multiple populations, or multiple classes, choice of the reference class depends on how relevant each class is, and how precise the statistical information. If all statistical information is sufficiently precise, then we try to use information about the most relevant class for which there is statistical information. The Bayesians part company here. Bayesians assume that there is precise statistical information about all classes; choice is a matter of maximum relevance. In Kyburg and Pollock, there may be no statistical information about the class of maximum relevance; of those about which there is information, there may be multiple classes of maximal relevance.

It may puzzle why we are interested in choosing among multiple statistcial sources when analogy reasons from a single source. Simply, analogy proceeds with a single source of imperfect match when there may be alternative sources and alternative matches. "Why this analogy?" may be closely related to "Why this class as reference class?"

The following table should help sort out the various approaches and the respects in which they should be compared.

Kyburg with 1-Projectibility	Clark	Russell
as few as one case suggests mass behavior	one case produces generalization	one case produces generalization
defeasibly	indefeasibly	indefeasibly
multiple classes	multiple determinations	multiple determinations
more specific classes are more relevant	no differential relevance	determinations have ideal relevance
can use multiple source cases	one source case	one source case

Kyburg	Subjective Bayes	Sampling
multiple cases or known frequencies suggest mass behavior	subjective estimates suggest mass behavior	multiple cases suggest mass behavior
defeasibly	without acceptance	at confidence
multiple classes	one ideally relevant source	no guidance on appropriate population
more specific classes are more relevant	condition on all relevant properties	no concept of relevance
can use multiple source cases	no source case	can use multiple source cases

2 Constructions.

2.1 Kyburg's Logical Foundations of Statistical Inference.

Kyburg's account is widely acknowledged among philosophers of science, but is quite complex; I assure the reader that this paper contains no more than the parts of theory required to follow the later discussion.

Kyburg formalizes reasoning about probabilities given statements about relative frequencies among classes.

Let me write

$$[\![A]\!] \text{ for } \{x : A(x)\}.$$

That means $[\![A \wedge B]\!]$ stands for $\{x : A(x) \wedge B(x)\}$. Also, let

$$\%([\![A]\!], [\![B]\!])$$

be the per cent of $[\![A]\!]$'s among $[\![B]\!]$'s.

Consider an example with three statistical sources of varying precision and relevance:

$$\%(\ [\![Mets\text{-}Game]\!], [\![Mets\text{-}Win]\!]\ ) = [.6, .7]$$
$$\%(\ [\![Mets\text{-}Game \wedge Night\text{-}Game]\!], [\![Mets\text{-}Win]\!]\ ) = [.5, .65]$$
$$\%(\ [\![(\lambda x).(x = Tuesday\text{'}s\text{-}Game)]\!], [\![Mets\text{-}Win]\!]\ ) = [0, 1]$$
$$Tuesday\text{'}s\text{-}Game \in [\![Mets\text{-}Game]\!]$$
$$Tuesday\text{'}s\text{-}Game \in [\![Night\text{-}Game]\!]$$
$$Tuesday\text{'}s\text{-}Game \in [\![(\lambda x).(x = Tuesday\text{'}s\text{-}Game)]\!]$$

apparently,

PROB (*Tuesday's-Game* $\in$ [|*Mets-Win*|]) = [.5, .65]

The reference class is [|*Mets-Game*$\wedge$ *Night-Game*|], the most specific class about which adequate statistics are known.

In the example, there are three candidates for the reference class (or candidate reference classes), for the query

PROB (*Tuesday's-Game* $\in$ [|*Mets-Win*|]) = ?

That is, there are three sets that satisfy the syntactic requirements required for projecting a *Mets-Win*. These sets are

[|*Mets-Game*|]
[|*Night-Game*|]
[|(λx).($x = Tuesday's\text{-}Game$)|]ie, $\{x : x = Tuesday's\text{-}Game\}$.

The last is the most specific class, but the relevant frequency is not well known among this class. Of the other two frequencies reported, there is conflict: *i.e.* [.6, .7] and [.5, .65] do not stand in the sub-interval relation, and the interval [.5, .65] is associated with the more specific class.

Formally, Kyburg takes an "inference structure" for PROB ($x \in V$) to be the collection:

$< x, A, V, [p, q] >$

> (Note I am suppressing Quinean quotes, but note that (x) is the universal quantifier in the object language, and the universal quantifier in the meta-language is spelled out, *for all* x; similarly, $\rightarrow$ is the material conditional in the object language, and "*if* ... *then* ..." is the material conditional in the meta-language).

when it is known that

$x \in A$
$\%(A, V) = [p, q]$
A MAY-PROJECT V,

7

where the last requirement is that A is a candidate reference class for V, *i.e.* A may project V.

(Also note that the first two sentences should be properly rendered
$$\ulcorner x \in A \urcorner \in \text{KBASE}$$
$$\ulcorner \%(A,V) = [p,q] \urcorner \in \text{KBASE}$$
to make them metalinguistic assertions, like the third sentence;
$$\ulcorner A \urcorner \text{ MAY-PROJECT } \ulcorner V \urcorner ;$$
I will resist this much propriety).

I will call V the target property and x the target individual. There are lots of classes that will not be candidate reference classes for V. For instance,

$$V \cup \{x\}$$

is clearly not a candidate reference class in general. This restriction is motivated by the same intuitions that lead to restrictions on what similarities may support analogy. We might not want $[\|Bob\text{-}Hope\text{-}did\text{-}not\text{-}attend\|]$ to be a candidate reference class for $[\|Mets\text{-}Win\|]$. Similarly, we might not take Russell's P, the property that a and b share in virtue of which an analogy is made, to be the property of not being identical to Bob Hope.

It is an axiom of Kyburg's system that

> *for all* A, B: *if* $\left(A \text{ MAY-PROJECT } V \right)$
> *and*
> $\left(B \text{ MAY-PROJECT } V \right)$
> *then* $\left(A \cap B \right) \text{ MAY-PROJECT } V$.

Given a collection of inference structures, $\{IS_1, IS_2, \ldots\}$, there will be relations of "domination" among them. In the example above,

DOMINATES $\Big($
$\quad < Tuesday's\text{-}Game,$
$\qquad [\|Mets\text{-}Game \wedge Night\text{-}Game\|],$
$\qquad [\|Mets\text{-}Win\|],$
$\qquad [.5, .65] > ,$
$\quad < Tuesday's\text{-}Game,$

8

$$[|Mets\text{-}Game|],$$
$$[|Mets\text{-}Win|],$$
$$[.6, .7] >$$
$$).$$

Let

$$\text{INF-STRUCTS}_{x,V}$$

be the set of inference structures for PROB $(x \in V)$; there will be one inference structure per (relevant) statistical source. Also, let

$$\text{REFCLASS}_{x,V}$$

be the reference class for PROB $(x \in V)$. When x and V are understood, we drop the indexes.

It is a theorem that

exists p, q *s.t.*
    $< x, \text{REFCLASS}, V, [p, q]> \in \text{INF-STRUCTS},$
    and
    not (
        exists IS' *s.t.*
            $IS' \in \text{INF-STRUCTS}$
            and
        DOMINATES $(IS', < x, \text{REFCLASS}, V, [p, q]>)$);

in other words, the inference structure that lists the reference class is undominated in the set of inference structures. Call the inference structure that lists the reference class the "reference inference structure." Say that two inference structures disagree when their intervals don't nest, that is,

DISAGREES (
    $< x, A, V, [p, q] >,$
    $< x', A', V', [p', q'] >$
)
  iff

$$\neg([p,q] \subset [p',q'])$$
$$and$$
$$\neg([p',q'] \subset [p,q]) \ .$$

Let I-REFCLASS(S) be the reference inference structure for a particular set of inference structures, S, which may not be the same as INF-STRUCTS. Note that I-REFCLASS(INF-STRUCTS) is REFCLASS.

The following two theorems for Kyburg's system establish how much of Clark's monotonicity of conflict resolution we will believe. The first says that if we remove a statistical source that disagrees with our ultimate conclusion, then the ultimate conclusion remains the same. The second says that if we add a statistical source that does not disagree with our ultimate conclusion, then either the conclusion is unchanged, or the new source becomes the new conclusion.

if
    $(IS_1 \in$ INF-STRUCTS$)$
    and
    DISAGREES $(IS_1,$ REFCLASS$)$
then
    not (
    DISAGREES (I-REFCLASS(INF-STRUCTS $— IS_1$), REFCLASS$)$
).

And

if
    not ($IS_2 \in$ INF-STRUCTS$)$
    and
    not (DISAGREES $(IS_2,$ I-REFCLASS(INF-STRUCTS$)))$
then
    I-REFCLASS((INF-STRUCTS $\cup IS_1$))
            $=$ I-REFCLASS(INF-STRUCTS$)$
            $=$ REFCLASS
    or
    I-REFCLASS((INF-STRUCTS $\cup IS_1$))
            $= IS_2$.

Incidentally, this system has been successfully implemented and studied as a system of evidential reasoning [Loui86,88]. These papers also discuss in more detail the combinations that are appropriate in situations in which there is more

than one undominated class. Combination of inference structures would take us too far afield, and there is no analogue in analogical reasoning. It suffices to say that here is another facet of statistical reasoning that has not yet been exploited in case-based reasoning (though Clark's attempts are a move in this direction).

2.2 Reconstruction of Russell's Inference.

Armed with this much understanding of Kyburg's system, we can now state Russell's and Clark's analogies as special cases of inferences with this system.

The point is that analogical inferences can be justified and understood in terms of an existing and accepted system of statistical reasoning.

The constructions for Clark's inferences are more interesting, so do not be dismayed by the triviality of claim 1.

> Claim 1. Russell's analogies can be understood as Kyburgian statistical inference given an example from the most specific candidate reference class, and given knowledge that the relative frequency of the target property is extreme in this class.

$P(x, y)$ DETERMINES $Q(x, z)$
$P(a, y) \wedge Q(a, z)$
$\underline{P(b, y)}$
thus,
$Q(b, z)$

becomes

1.1.1. *for all* y, z:
  $[|(\lambda x).P(x, y)|]$ MAY-PROJECT $[|(\lambda x).Q(x, z)|]$
      and
  for all A, t:
      if
          $(t \in [|(\lambda x).P(x, y)|])$
          and $(t \in A)$
          and A MAY-PROJECT $[|(\lambda x).Q(x, z)|]$
      then
          $[|(\lambda x).P(x, y)|] \subseteq A$

1.1.2. $(y)(z).$
  $\%(\; [|(\lambda x).P(x, y)|], \; [|(\lambda x).Q(x, z)|] \;) = [1, 1]$
  $\vee$

$$\%(\ [\|(\lambda x).P(x,y)\|],\ [\|(\lambda x).Q(x,z)\|]\ ) = [0,\ 0]$$

1.2. $a \in [\|(\lambda x).P(x,y_1) \wedge (\lambda x).Q(x,z_1)\|]$

1.3. $\underline{b \in [\|(\lambda x).P(x,y_1)\|]}$

thus,
1.4. $\textsc{prob}(\ b \in [\|(\lambda x).Q(x,z_1)\|]\ ) = [1,\ 1]$

(the numbering is suggestive of the mapping between the premises).

Proof of Claim 1. Sufficiency. Without assuming details of Kyburg's system, I can only sketch the proof. The critical observations are that

o 1.2 and 1.1.2 fix the % of Q's among P's.

o 1.3 says that this leads to an inference structure for the probability that b is a Q.

o There may be other disagreeing inference structures, but 1.1. guarantess that they are dominated.

Note that here as in Russell's account, the source case makes the inference about the target indefeasible. We have taken the probabilities to be extremes. If b is known not to be $(\lambda x).Q(x, z_1)$, then that is an inconsistency, not a defeater of the analogical inference.

1.1.1 can be weakened by taking just its first conjunct. 1.4 would still be a conclusion if there are no known competing inference structures for b's Q-ness or $\neg Q$-ness. But note that Russell requires that the predicate P reflect "all the information relevant to the query, ...for example, ...all the factors that might affect the language a person speaks (nationality, country of residence, parents' language, ...and so on) ...." This is reflected in the requirement in (1.1.1) that $[\|P(x,y)\|]$ be the most specific candidate reference class to which any individual will be known to belong. We require any individual, not just the target individual, because Russell's determination relation says nothing about a particular individual. Usually, $[\|(\lambda x).(x = t)\|]$ is taken to be a candidate reference class in Kyburg's system, for all queries, but this almost surely will not be the case here. There has to be some reason why, when we do analogy, we are so certain that the inference from the putative determining case will not be defeated by an inference from an even more similar determining case. In Russell, and here, this certainty is simply imposed by fiat. The assumption of certainty is relaxed in Clark's study.

12

Russell worries about what should happen when Louis speaks two languages, or has two nationalities, and he alters the definition of a determination several times to account for this case. Here, too, there would be complications, but I will ignore them and focus only on total determinations. Accounts of his other determination relations could be given, but would just be distracting.

2.3 Reconstruction of Clark's Inference.

Claim 2. Clark's analogies can be understood as Kyburgian statistical inference in the presence of multiple statistical sources, when there is knowledge about how conflicts were decided for the source case.

It cannot be understood as using the single case to determine the frequency of the target property among the intersection of candidate reference classes. This is because even if the single case could determine the frequency, and we could project from this single case, we could not explain Clark's insistence on the monotonicity of conflict resolution.

$A_1(x)$ INDICATES $\neg Q(x)$
$A_2(x)$ INDICATES $Q(x)$
$A_3(x)$ INDICATES $Q(x)$
$A_4(x)$ INDICATES $\neg Q(x)$
$Q(a)$
$A_1(a) \wedge A_2(a) \wedge A_3(a) \wedge A_4(a)$
$\underline{A_1(b) \wedge A_2(b) \wedge A_3(b) \wedge A_4(b)}$
thus,
$Q(b)$

becomes

2.1.1. $\%(\, [|A_1|], \, [|Q|] \,) = [0, \epsilon]$
2.1.2. $[|A_1|]$ MAY-PROJECT $[|Q|]$

2.2.1. $\%(\, [|A_2|], \, [|Q|] \,) = [1 - \epsilon, 1]$
2.2.2. $[|A_2|]$ MAY-PROJECT $[|Q|]$

2.3.1. $\%(\, [|A_3|], \, [|Q|] \,) = [1 - \epsilon, 1]$
2.3.2. $[|A_3|]$ MAY-PROJECT $[|Q|]$

2.4.1. $\%(\, [|A_4|], \, [|Q|] \,) = [0, \epsilon]$
2.4.2. $[|A_4|]$ MAY-PROJECT $[|Q|]$

2.5. $a \in [|A_1 \wedge A_2 \wedge A_3 \wedge A_4|]$

2.6. $Q(a)$ is acceptable on $2.1 - 2.5$;
$\quad$ *i.e.* PROB $_{2.1-2.5}(\ a \in [|Q|]\ ) \geq 1 - \epsilon$

2.7. $\underline{b \in [|A_1 \wedge A_2 \wedge A_3 \wedge A_4|]}$

thus,
2.8. PROB $(b \in [|Q|]) \geq 1 - \epsilon$

Proof of Claim 2. Sufficiency. The crucial observations are

o There are as many inference structures for $b \in [|Q|]$, as for $a \in [|Q|]$, given $2.1 - 2.5$.

o There may be other reasons for $a \in [|Q|]$, but on the force of the inference structures that $2.1 - 2.5$ support, PROB $(a \in [|Q|])$ is high.

o So some inference structure with reference class Y, with $\%(Y, [|Q|])$ high, dominates other inference structures (whether this reference inference structure is based on 2.2, 2.3, or some combination of them).

o If this inference structure is dominating for $a \in [|Q|]$, then it is dominating for $b \in [|Q|]$.

What is more interesting than sufficiency is the insufficiency of a weakened 2.6. If 2.6 is weakened, so that $Q(a)$ is simply acceptable, then this doesn't guarantee 2.8. $Q(a)$ may be acceptable because of some other inference structure for $a \in [|Q|]$, and there may be no analogous inference structure for $b \in [|Q|]$.

Note that monotonicity of conflict resolution is achieved through the theorems on how I-REFCLASS behaves under unions and subtractions from INF-STRUCTS. If inference structures dominate each other in such a way that warrants concluding PROB $(b \in [|Q|])$ exceeds some threshold, then adding yet another inference structure that agrees with that conclusion cannot force the threshold lower. This corresponds to Clark performing the analogy from a source case that has fewer factors indicating the ultimate conclusion than the target case, but just as many counter-indicating factors. Also, if we remove from the set of inference structures an inference structure that disagrees with the conclusion about PROB $(b \in [|Q|])$, that too cannot lower the threshold. This corresponds to Clark performing the analogy from a source case that has as many indicating factors for the ultimate conclusion, but even more counter-indicating factors.

It is tempting, but incorrect to take the translation to be

3.1.1. $\%(\; [\|A_1\|], \; [\|Q\|] \;) = [0, \; \epsilon]$
3.1.2. $[\|A_1\|]$ MAY-PROJECT $[\|Q\|]$

3.2.1. $\%(\; [\|A_2\|], \; [\|Q\|] \;) = [1 - \epsilon, \; 1]$
3.2.2. $[\|A_2\|]$ MAY-PROJECT $[\|Q\|]$

3.3.1. $\%(\; [\|A_3\|], \; [\|Q\|] \;) = [1 - \epsilon, \; 1]$
3.3.2. $[\|A_3\|]$ MAY-PROJECT $[\|Q\|]$

3.4.1. $\%(\; [\|A_4\|], \; [\|Q\|] \;) = [0, \; \epsilon]$
3.4.2. $[\|A_4\|]$ MAY-PROJECT $[\|Q\|]$

3.5. $\%(\; [\|A_1 \wedge A_2 \wedge A_3 \wedge A_4\|],$
$\qquad [\|Q\|] \;) = [1 - \epsilon, \; 1]$
$\quad \vee$
$\quad \%(\; [\|A_1 \wedge A_2 \wedge A_3 \wedge A_4\|],$
$\qquad [\|Q\|] \;) = [0, \; \epsilon]$

3.6. $a \in [\|A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge Q\|]$

3.7. $b \in [\|A_1 \wedge A_2 \wedge A_3 \wedge A_4\|]$

thus,
3.8. PROB $(b \in [\|Q\|]) \geq 1 - \epsilon$.

First, the single case reported in (3.6) does not establish which of the disjuncts in (3.5) is true. Suppose, though, on the basis of (3.6), one of the disjuncts could be made so probable that it is acceptable, through the observation of the target case. This requires an additional theorem about sampling: that most of the candidate reference classes for some property y that are known to be mostly y or mostly $\neg y$, can be decided (with high probability) by looking at a single case. This is a theorem for appropriate ϵ and δ.

3.5.1. (y).
$\qquad \%($
$\qquad \{ \; x \; : \; (\%(x,y) = [0, \; \epsilon] \vee \%(x, y) = [1 - \epsilon, \; 1])$
$\qquad\qquad \wedge \text{ sample}(x, y) = \; <1, \; 1> \; \},$
$\qquad \{ \; x \; : \; \%(x, y) = [1 - \epsilon, \; 1] \; \}$
$\qquad) \geq 1 - \delta$.

15

(where, by sample$(P, Q) = <s, r>$, we mean that sP's were sampled, of which r were Q's and $s - r$ were $\neg Q$'s).

Additional knowledge about particular x's and y's, or for particular sampling procedures, might make δ even smaller, for given ϵ.

When we have a situation like (3.5), and the δ for which (3.5.1) is a theorem is small enough to suit our needs (*i.e.* $1 - \delta$ is above the threshold of acceptability), then let us write that Q is single-case-projectible from $A_1 \wedge \ldots \wedge A_4$; that is:

$$[|A_1 \wedge \ldots \wedge A_4|]_{\text{1-PROJECTS}} [|Q|].$$

Then (3.6) decides which disjunct in (3.5) is the right one; the mass of $A_1 \wedge \ldots \wedge A_4$ is either Q or $\neg Q$, and since a is a Q, this decides that the bulk of $A_1 \wedge \ldots \wedge A_4$ is Q with high probability. So $Q(b)$ could be inferred.

But the monotonicity of conflict resolution would be violated for the following reason. If

$$\%(\ [|A_1 \wedge A_2 \wedge A_3 \wedge A_4|],\ [|Q|]\ )$$
$$= [1 - \epsilon,\ 1]$$
$$\wedge$$
$$\%(\ [|A_1|],\ [|Q|]\ ) = [0,\ \epsilon],$$

it does not follow that

$$\%(\ [|A_2 \wedge A_3 \wedge A_4|],\ [|Q|]\ )$$

can be bounded. In particular, it is not necessarily close to one. If $b \in [|A_1 \wedge A_2 \wedge A_3 \wedge A_4|]$, indeed, $Q(b)$ can be projected. But if (3.7) is altered, if b is only known to be in $[|A_2 \wedge A_3 \wedge A_4|]$, no projection is possible. It could be that most races in which there are Porsches are not close races, most races in which there are Ferraris are not close races, but races in which there are both Porsches and Ferraris are often close races.

So in this formulation of the inference, we cannot weaken the number of counter-indicating factors for the source case. It may be that the interaction of two counter-indicating factors produces a joint indicating factor. Why could we weaken it in the first reduction of Clark? In that reduction, the combination of two counter-indicating inference structures can only counter-indicate. So the counter-indicators for the source could be weakened without disturbing the conclusion. Combination of inference structures is related to the cross product

of two sets, while joint effects are related to the intersection of two sets; therein lies the difference.

Anyway, we do allow the other half of Clark's monotonicity desideratum. We can know one more indicating factor about b, the target, and continue to draw the conclusion. Add

3.7.1. $b \in [|A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge A_5|]$

3.7.2. $\%([|A_5|], [|Q|]) = [1 - \epsilon, 1]$

3.7.3. $[|A_5|]$ MAY-PROJBOT $[|Q|]$

and (3.8) is still a conclusion in Kyburg's system.

3 Discussion.

3.1 Considerations on Clark's Defeasible Analogical Reasoning.

It is fallacious to propose a reduction of one inference system as a special case of another, then criticize it because it appears to make assumptions in the language of the reducing system. It is fallacious because there is no guarantee that the reduction is the only possible, and there is no guarantee that ensuing disputes are not defects of the reducing system. This is what the Bayesians are sometimes guilty of doing when they assail Dempsterian inference.

I will not do this. But there are some claims about Clark's analogical inferences I would like to make, and I take the reductions to be merely suggestive of the truth of those claims.

The reason I have belabored the second formulation of Clark-like inference is that it is an epistemologically more satisfying account of analogy. It does not capture both halves of Clark's monotonicity of conflict resolution. But it is simpler to use the second reduction and not insist on the ability to project from a source to a target when more relevant things are known about the source than about the target (namely, that the source has additional counter-indicating factors of the ultimate conclusion).

The latter account requires trivial background knowledge; it requires almost no assumptions in order to reconstruct Clark's inferences. Domination of inference structures based on specificity of candidate reference classes does all the work. There are only two specializations. The first is that when two factors both indicate a conclusion, they do so with frequency intervals that nest. That is, if A_1 indicates Q, and A_2 indicates Q, then $\%(A_1, Q)$ and $\%(A_2, Q)$ should stand in the (possibly improper) sub-interval relation. I guaranteed this above by taking all the intervals for indicating factors to be $[1 - \epsilon, 1]$ (and all the intervals for counter-indicating factors to be $[0, \epsilon]$). This could have been done

in a more general way: by taking all the intervals for indicators to be anchored on the right at 1 (and taking all the intervals for counter-indicators to be anchored on the left at 0). The second specialization is in (3.5.1), which says that observation of a single case determines the statistics to be on one extreme or the other, hence allows projection from that single case. (3.5.1) is a very natural assumption about sampling, for the appropriate δ.

In the former account, which preserves both parts of the monotonicity of conflict resolution, the source case had to be treated in a special way. It was not enough to say in (3.6) that $Q(a)$ was in fact accepted, for instance, by some fortuitous observation. It had to be that $Q(a)$ was acceptable, on the basis of a's having properties A_1, A_2, A_3, and A_4. $Q(a)$ could then be accepted by inference. Or perhaps that inference was unnecessary because $Q(a)$ had already been observed; it wouldn't matter which. But it is not enough to say that *Sand-Nearby*, and *Unfavorable-Environment*, and *Sand*, among others, all co-occurred at *well*$_1$. It must be that their co-occurrence was not accidental, that this was a special kind of co-occurrence that determines how conflicting arguments of exactly the kind that were involved are to be resolved, for all future cases, including *well*$_2$. It must be that the co-occurrence of *Louis's Nationality* and his *Language* is not spurious; rather, that *Louis's Nationality* and *Louis's Language*'s co-occur representatively.

3.2 Defeasible Determination and Defeasible Representativeness.

In the case of Russell-like analogies, since there is no defeasibility, the determination relation dictates that any source example will relate Nationality and Language in a representative way.

In Clark-like analogies, the stipulation of this representativeness is the intuition that drives the design of the inference system. Any source example that involves certain factors and their resolution is representative of all such resolutions.

Clark puzzles over how to generalize the formalism so that cases with the same indicators and counter-indicators could have resolved differently in past, arriving at different ultimate conclusions. He has taken the first step of introducing defeasibility into determinations. Now a factor, such as *French-Nationality*, only appears to determine whether *French-Language* or $\neg$*French-Language*; it is a prima facie determiner. It defeasibly determines, and the ultimate conclusion about whether *French-Language* or $\neg$*French-Language* in a particular case depends on what other defeasible determinations play mitigating roles, such as $\neg$*French-Nationality-Parents* defeasibly determining whether *French-Language* or $\neg$*French-Language*.

Clark puzzles over the next step, introducing defeasibility into representativeness. How should conflict resolution be done defeasibly?

We may consider the inference

$A_1(\mathbf{x})$ INDICATES $\neg Q(x)$
$A_2(\mathbf{x})$ INDICATES $Q(x)$
$A_3(\mathbf{x})$ INDICATES $Q(x)$
$A_4(\mathbf{x})$ INDICATES $\neg Q(x)$
$A_1(a_1) \wedge A_2(a_1) \wedge A_3(a_1) \wedge A_4(a_1) \wedge Q(a_1)$
$A_1(a_2) \wedge A_2(a_2) \wedge A_3(a_2) \wedge A_4(a_2) \wedge Q(a_2)$
$A_1(a_3) \wedge A_2(a_3) \wedge A_3(a_3) \wedge A_4(a_3) \wedge \neg Q(a_3)$
$\underline{A_1(b) \wedge A_2(b) \wedge A_3(b) \wedge A_4(b)}$
thus,
$Q(b)$?

Clark mentions that a "majority verdict" could be taken, in this case, two cases, a_1 and a_2, are in favor of Q, while the single case, a_3, is opposed.

Are we willing to choose Q over $\neg Q$ for b when no overwhelming number of past cases are examples of Q? I think we are not. If the number of cases for Q were overwhelming, say 10 against 1, then perhaps we should project $Q(b)$. Otherwise, by my lights, there is no sense to the inference, even if possibly construed as some strong form of analogy.

The only sense I can make of such an inference is that we are accumulating instances of the class $[[A_1 \wedge A_2 \wedge A_3 \wedge A_4]]$, and observing the relative frequency among them that is is $[[Q]]$. Our willingness to project from this class based on a small number of cases includes those times when we have a sample of 1 among 1, or 2 among 2, perhaps even 9 among 10; but not 2 among 3, or 5 among 11.

Fully defeasible analogical reasoning is apparently just defeasible reasoning about the reference class, with the right to project from certain small samples.

3.3 Defeasible Reasoning with Specificity.

Oddly, Clark makes no use of the idea of specificity, which pervades defeasible statistical reasoning, and defeasible reasoning in general [Loui87]. If A_1 indicates Q, while $A_1 \wedge A_2$ counter-indicates Q, the latter should supercede the former for a case t, *s.t.* $A_1(t) \wedge A_2(t)$. This is not a particularly egregious oversight, since Clark's whole philosophy is to make no a priori distinctions among determining factors; he wants to resolve all conflicts by looking for source cases that manifest these factors and seeing how their conflicts were resolved. But he misses an opportunity by not seeing the source case as a more specific reason than the indication relation.

He could be faced with the case above, where A_2 and A_3 are both indicators of Q, and A_1 and A_4 are both counter-indicators of Q. There are two potential sources: a_1 and a_2. a_1 is an A_1 and an A_2, as well as a $\neg Q$. a_2 is an A_1 and an A_2 and an A_4, as well as a Q. In this case, though one source argues for Q and the other for $\neg Q$, so that there is no majority verdict, taking specificity seriously

would demand the (defeasible) conclusion $Q(b)$. a_2 is just more specific a source than a_1.

3.4 Case-Based Generalization versus Case-Based Projection.

The difference between generalization and projection for arbitrary cases is that the projection is defeasible. It is one thing to infer

P DETERMINES Q
$$\frac{P(a) \wedge Q(a)}{thus,}$$
$(x). \ P(x) \rightarrow Q(x)$

and quite another to say

P DETERMINES Q
$$\frac{P(a) \wedge Q(a)}{thus,}$$
for all x: *if* $P(x)$ is known,
     then $Q(x)$ is a justified defeasible conclusion .

The two are not the same because when $\vdash$ is taken to be non-monotonic, the deduction theorem disappears. The former conclusion indefeasibly yields Q from P. The latter admits defeaters.

What really seems to be desired is the latter.

"Case-based generalization" is a misnomer. Clark and Russell understand the logic of analogy to be monotonic. I think this is a mistake.

4 Control of Statistical Reasoning Driven by Few Cases.

The possibility of case-based reasoning suggests new control strategies for statistical reasoning programs. Already there is a need for studying control issues in programs that pore over their database in response to a query, to construct samples, from which to do reasoning about reference classes (*e.g.* the program described in [Loui88]). It makes little sense to continue to extend sample sizes beyond 10, or 20, in commonsense reasoning. There is hardly a difference in the intervals that result from a sample of 7 out of 10, and a sample of 14 out of 20.

If I had time to examine 20 cases, I would spot at 7 out of 10 for one class, and spend the rest of my counting time in its subclasses.

A meta-level utility analysis of the expected value of continued examination of cases is required, no doubt (see for example, [Russell&Wefald88]).

Prior to doing such an analysis, we should identify what choices make sense. These choices are determined by what are our current best arguments for projecting some property, and what counter-arguments, if constructible, would be effective rebuttals. Our theory of analogy, or of more general projection from few cases, is what determines where to seek counter-arguments.

Consider projecting from the single case:

$Porsche(944) \wedge Powerful(944)$
$\underline{Porsche(924)}$
apparently,
$Powerful(924)$

There are a few ways in which search can now be directed. A counter-argument to this projection could produce counter-examples to the co-occurrence, *e.g.*

$Porsche(356) \wedge \neg Powerful(356)$

which dilutes the relevant statement of relative frequency. Or it could produce a property that distinguishes putative target and source, *e.g.*

$VW\text{-}Project(924) \wedge \neg VW\text{-}Project(944)$

together with a statement that this property is a counter-indicating factor, that is, that it indeed leads to an interfering inference structure

$VW\text{-}Project(Bug) \wedge \neg Powerful(Bug)$.

This amounts to an appeal to the target's membership in a more specific (*i.e.* more relevant) class that is apparently counter-indicating.

A third way to attack the argument is to attack the determination relation in virtue of which the single case can be projected. For instance, if one could produce

$$Ford(Mustang) \land Powerful(Mustang)$$
$$Ford(Escort) \land \neg Powerful(Escort)$$

it would be contentious whether

for all x, y:
if $x \in Car\text{-}Manufacturers$
and $y \in Performance\text{-}Features$
then x 1-PROJECTS y .

This amounts to attacking the inference that led to the 1-projectibility relation between *Porsche* and *Powerful*, that is being assumed (in the real world, we probably already know this relation, so the *Ford* examples are not going to sway our opinion).

The *Ford* example is prima facie evidence that there is a *Car-Manufacturer* class that is inhomogeneous in some *Performance-Feature* class. This example could be countered by identifying *Porsche* in a subclass, *e.g. European-Marques*, where the *Ford* example is excluded from this subclass. And there should be prima facie evidence that there is homogeneity of members of this class with respect to kinds of performance features; there should be reason to believe that a *Performance-Feature* could be 1-projected from a *European-Car-Marque*:

$$\%([|Saab|], [|Handles|]) = [1 - \epsilon, 1]$$
$$\%([|Yugo|], [|Handles|]) = [0, \epsilon]$$
$$\%([|Lotus|], [|Brakes\text{-}Well|]) = [1 - \epsilon, 1]$$
$$\%([|Daimler\text{-}Benz|], [|Slow|]) = [0, \epsilon]$$

This is a lot of background knowledge, and in the presence of such knowledge we might as well assume that we also know

$$\%([|Porsche|], [|Powerful|]) .$$

But it should be clear how the logic could direct the dialectic. If any of these counter-arguments could be produced, search could next be directed toward finding a reinstating argument, and so forth.

22

Investigations of control of this kind of reasoning would also benefit control of purely qualitative defeasible inference, such as [Loui87] and [Pollock87].

How precisely to make control choices appropriate to this kind of reasoning is the subject of a more ambitious invesitigation. For our purposes, it is enough to recognize this kind of reasoning, and to recognize the relation between control of choices and the study of how to undermine analogical inferences.

5 Conclusion.

Modern philosophers of science such as Quine and Ullian [Quine&Ullian70] and Kyburg [Kyburg61] hold that analogy is an uninteresting special case of induction. AI authors have balked at this, preferring to view analogy as a more interesting species of deductive reasoning, in Russell's case, or as the resolution of conflicting defeasible arguments, in Clark's case. I present middle ground. Analogy's problems are best sorted with our most expressive language and machinery for inductive statistical reasoning. Case-based reasoning reveals itself as the guide to dialectical maneuvers in this statistical reasoning.

6 Bibliography.

Clark, P. "Representing arguments as background knowledge for the justification of case-based inferences," AAAI Workshop on Case-Based Reasoning, St. Paul, 1988.

Davies, T. and Russell, S. "A logical approach to reasoning by analogy," IJCAI, 1987.

Gentner, D. "Structure Mapping," *Cognitive Science 7*, 1983.

Kedar-Cabelli, S. "Purpose-Directed Analogy," Cognitive Science Society Meeting, 1985.

Kyburg, H. *Probability and the Logic of Rational Belief*, Wesleyan, 1961.

Kyburg, H. *Logical Foundations of Statistical Inference*, Reidel, 1974.

Kyburg, H. "The Reference Class," *Philosophy of Science 50*, 1982.

Loui, R. "Computing reference classes," AAAI Workshop on Uncertainty, Philadelphia, 1986.

Loui, R. "Defeat among arguments", *Computational Intelligence 3*, 1987.

Loui, R. "Evidential reasoning in a network usage prediction testbed," AAAI Workshop on Uncertainty, St. Paul, 1988.

Pollock, J. "A theory of direct inference," *Theory and Decision 16*, 1983.

Pollock, J. "Foundations for direct inference," *Theory and Decision 17*, 1984.

Pollock, J. "Defeasible reasoning," *Cognitive Science 12*, 1987.

Quine, W.V.O. and Ullian, J. *The Web of Belief*, Random House, 1970.

Russell, S. Analogical and Inductive Reasoning. Ph.D. dissertation, Stanford University, Dept. of Computer Science. Also STAN-CS-87-1150, 1987.

Russell, S. and Wefald, E. "A Meta-Greedy Game-Playing Algorithm," draft, Computer Science Division, UC Berkeley, 1988.

Weitzenfeld, J. "Valid Reasoning by Analogy," *Philosophy of Science 51*, 1984.

Winston, P. "Learning and Reasoning by Analogy," *CACM 23*, 1980.