Caricatures: Generating Models of Dominant Behavior

Olivier Raiman, Brian C. Williams*
Xerox Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, CA 94304 USA

Abstract

When analyzing most physical phenomena the complexity of the corresponding equations rapidly becomes such that there are no general methods to derive exact solutions. A solution prevalent in acid-base chemistry is to construct a simplified solution that preserves only the dominant behaviors. Since what dominates varies, for example, depending on the strength of the acid or its initial concentration, the chemist divides behavior into a patchwork of simpler subregimes that reflect these variations. The success of such an approach hinges upon the careful identification of the simplifying assumptions (e.g., the acid is strong) which induce the partition. What is most striking is that these assumptions appear to arise prior to thoughts about how the models are to be used — modelling is an emergent process.

To identify dominant regimes we exploit the metaphor of a *caricature* — an exageration of an equation's prominent features — to generate the requisite simplifying assumptions. Generating these assumptions, the boundaries of the partition, and the simplified equations for each regime, draws heavily upon our earlier work on qualitative algebraic and order of magnitude reasoning. The resulting process, called *caricatural modeling*, is sufficient to replicate a broad set of examples from acid-base chemistry.

Introduction

Along with others we have argued [0, 0, 0, 0] that the model generation/selection task is ultimately driven by the phenomena of interest, as dictated by the problem being solved. Thus, we found it striking, when examining analytical chemistry texts[0], that much time is devoted to teaching modeling skills at the start — with no mention of how these models are to be used. We believe this is an instance of a novel and pervasive facet of modeling that requires explanation.

*The ordering of authors is incidental.

We claim that modeling is often an emergent phenomena — there exists a distinct notion of interesting phenomena, which is not contingent on the task being performed. The argument goes as follows: People are extremely inventive; they are quite good at making use out of just about anything they understand, whether it be the design of a mechanical device or a chemical synthesis.[1] The difficult issue then is to come up with models that, while accurate, are sufficiently simple to be intuitively grasped. Thus, at least for invention, *the problem of constructing simple, but accurate models may precede use.*

We observe that emergent models achieve simplicity, by highlighting dominant behaviors, and carving a system into a patchwork of regimes where different behaviors dominate. Finally, we claim this patchwork emerges as *caricatures* of the system, by reinforcing its prominent features. We present a domain independent approach to generating emergent models, called *caricatural modeling*, and demonstrate it in the context of acid-base chemistry.

An example from chemistry

Consider equilibrium behavior of a simple reaction — the dilution of acid molecules, AH, into water. The dilution is characterized by the reactions $H_2O \rightleftharpoons H^+ + OH^-$ and $AH \rightleftharpoons H^+ + A^-$, and its equilibrium state is governed by:

$$\begin{array}{llcl}
\text{(I1) Charge balance:} & h^+ & = & oh^- + a^-, \\
\text{(I2) Mass balance:} & C_a & = & ah + a^-, \\
\text{(I3) Water equilibrium:} & K_w & = & h^+ oh^-, \\
\text{(I4) Acid equilibrium:} & K_a ah & = & h^+ a^-,
\end{array}$$

where ah, a^-, h^+, h_2o, and oh^- denote the concentration[2] at equilibrium of the species AH, A^-, H^+, H_2O, and OH^-, respectively, K_w and K_a are equilibrium constants for the water and acid ionization,

[1] The approach of focussing foremost on making interactions tractable during invention we call interaction-based design.[0]

[2] Although the concentration of species S is traditionally denoted $[S]$, this conflicts with the use of [ ], within qualitative reasoning, to denote a quantity's sign.

and C_a denotes the initial concentration of AH. Coming up with these equations is straightforward; the most reasoning intensive step is to solve for the equilibrium concentrations.

Bypassing brute force through a patchwork of dominant behaviors

To derive, for example, the concentration of H^+ ions we could use a brute force approach to solving the system of nonlinear equations. Eliminating ah, a^-, and oh^- in the four equations yields the following *equilibrium concentration equation*:

$$h^{+^3} + K_a h^{+^2} - (K_a C_a + K_w)h^+ = K_a K_w.$$

The derivation of this equation is used in chemistry texts to make clear the need to avoid brute force. Even for this simple case, the concentration equation is a third degree polynomial in h^+, making it difficult to solve (solving this equation involves computing the real roots of the left hand side). For more complicated cases, such as polyprotic acids — acids with more than one replaceable hydrogen ion H^+, such as H_3PO_4 — the degree of such equations increases with the number of replaceable ions. For example, H_3PO_4 has three replaceable ions and results in a concentration equation of degree five. Of course, there are no general solutions to algebraic equations of degree five or higher (by Galois). Likewise, deriving each concentration equation involves solving a system of non-linear equations, which can be quite computation intensive. Thus, applying brute force reaches a dead end for all but the simplest cases.

Instead, a chemist is taught to proceed as follows (taken from [0], chapter 5). First, having introduced equations I1–I4, governing the reaction's equilibrium, the chemist guesses several interesting simplifying assumptions about what the dominant species may be:

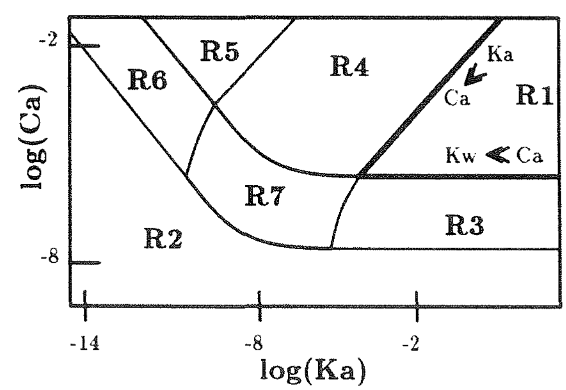
A1: The acid is weak ($a^- \ll C_a$).
A2: The acid is strong ($ah \ll C_a$).
A3: The soln. is essentially neutral ($a^- \ll h^+$).
A4: The soln. is strongly acidic ($oh^- \ll h^+$).

Combining, for example, assumptions A2 and A4 and applying them to the charge and mass balance equations (I1,I2) produces $h^+ \approx a^-$ (I1') and $C_a \approx a^-$ (I2'). Solving for h^+ results in $h^+ \approx C_a$, a far simpler result than produced through brute force.

Applying other combinations of assumptions in a similar manner produces:

Assump	Rgn	Simplified Concentration Eqn
A2,A4	R1	$h^+ \approx C_a$ (E1)
A3	R2	$h^{+^2} \approx K_w$.
A2	R3	$h^{+^2} - C_a h^+ \approx K_w$
A4	R4	$h^{+^2} + K_a h^+ \approx C_a K_a$
A1,A4	R5	$h^{+^2} \approx C_a K_a$
A1	R6	$h^{+^2} \approx C_a K_a + K_w$
none	R7	$h^{+^3} + K_a h^{+^2} - (K_a C_a + K_w)h^+ = K_a K_w$

The remaining step is to determine the *domain of validity* for each set of assumptions — the constraints that the assumptions impose on the givens, K_a, K_w and C_a. Returning to the pair of assumptions A2 and A4, from A2 ($ah \ll C_a$) we derive $C_a^2/K_a \ll C_a$ by substituting for ah and a^- using the simplified acid equilibrium (I4), acid concentration (E1) and mass balance (I2') equations. And from A4 ($oh^- \ll h^+$) we derive $K_w/C_a \ll C_a$ by substituting for oh^- and h^+ using the simplified acid concentration (E1) and water equilibrium (I3) equations. These two constraints define a region, R1, whose fringe corresponds to the two bold lines in the upper right corner of region diagram below (taken from [0], p. 75). The domains of validity R2–R7 for the remaining sets of assumptions partitions the reaction's behavior into simpler regimes according to the values of C_a and K_a. Given these results, the problem of identifying a solution's acidity for given values of C_a and K_a involves identifying the appropriate region and applying the corresponding simplified concentration equation. This paper demonstrates how to automate this style of reasoning.



For $K_w = 10^{-14}$ $a \ll b$ when $a/b < 0.1$

Elements of the chemist's expertise

What are the essential characteristics of the above example? The intractability of solving the initial equations is avoided from the start by replacing them with sets of equations which are simpler to solve, and are still relatively accurate renditions of the initial equations. An accurate, yet simpler rendition is achieved through a pervasive style of reasoning: first the complex behavior is partitioned into a set of regimes where subsets of the behavior dominate; then the equations are approximated by eliminating all but the dominant behaviors in that regime. The resulting equations are easily solved, circumventing the need for sophisticated mathematics.

The essential skill — which is deeply rooted in the chemist's know-how but poorly systematized — is the ability to identify the dominant regimes and their corresponding simplifying assumptions. We offer here an approach for identifying such regimes, which is sufficient to replicate a broad set of acid-base chemistry examples (taken from [0]), and draws heavily on our earlier work on qualitative algebraic reasoning involving hybrid qualitative/quantitative [0] and order of magnitude algebras [0]. This approach is based on the metaphor of a *caricature*, which provides clues to what are the interesting simplifying assumptions and the corresponding regimes. Its instantiation is a process we call *caricatural modeling*.

What is a caricature?

From a commonsense standpoint a caricature of an object is a description which exagerates prominent features, and eliminates insignificant features. For example caricatures of Richard Nixon reduce his face to little more than a nose with an exaggerated slope. Applying this concept to modeling, given a system of initial equations we construct a *caricature of this system* by exagerating one or more of the equation's prominent features. In this paper, we take "prominent" to mean that one term a of an equation E dominates another term b: $|a| > |b|$; that is, a is further from zero than b. We exagerate this feature by making $|a|$ much greater than $|b|$, thus making a dominant and b insignificant: $|a| > |b| \rightsquigarrow a \gg b$. We call this relation a *caricatural assumption*. By using this assumption to simplify E we produce a *caricatural equation*, which eliminates the insignificant features.

For example, given that all concentrations are positive, two prominent features of equation I2 ($C_a = ah + a^-$) are $|C_a| > |ah|$ and $|C_a| > |a^-|$ (note that all concentrations are positive). Exagerating $|C_a| > |ah|$ introduces the caricatural assumption $C_a \gg ah$, and allows I2 to be replaced by the caricatural equation $C_a \approx a^-$. This corresponds to the chemist's notion of a strong acid (i.e., essentially all AH dissociates). Conversely, exagerating $|C_a| > |a^-|$ introduces the assumption $C_a \gg a^-$, and produces the caricature $C_a \approx ah$, the chemist's notion of a weak acid (i.e., a negligible fraction of the acid AH dissociates).

Of course an alternative approach might take a quantity or subterm from *any* two equations and presume one dominates another. However, the number of potential assumptions would be prohibitively large. Instead the concept of caricature allows us to use existing features of the initial equations as clues to what relations are worth exaggerating. What is striking is that the restricted set generated through caricatures matches the simplifying assumptions introduced in a variety of acid-base chemistry examples.

There are two additional issues. First, more than one feature may be exagerated. For example, portraits of Nixon often exagerate both his nose and jowls. Likewise we might exagerate C_a relative to both ah and a^-. Second, exagerating a set of features may not always be consistent. For example, combining both caricatural assumptions with equation I2 allows us to conclude that $C_a \gg C_a$, which is inconsistent for positive C_a.

Following the example, after identifying the caricatural assumptions and generating caricatures of the initial equations, what remains is to solve for the simplified concentration equations and the domain of validity for the caricatural assumptions. As we elaborate in the next few sections, caricatural modeling involves 1) generating sets of caricatural assumptions, 2) deriving the caricatures of the initial equations, 3) solving for the simplified concentration equations, 4) deriving the domains of validity, and 5) recognizing inconsistent sets of assumptions. Space precludes a detailed presentation of the algebraic manipulations, described elsewhere in [0, 0]. Rather, our goal is to demonstrate how, by exploiting these algebraic manipulation techniques, caricatural modeling is able to replicate the chemist's tacit skills.

Generating caricatural assumptions

We begin by extracting the prominent features from each initial equation. We express both features and equations using the hybrid qualitative/quantitative algebra SR1 (an algebra combining signs and reals). Briefly, the domain of SR1 extends the reals to include signs (i.e., $\hat{+} \equiv (0, \inf)$, $\hat{-} \equiv (-\inf, 0)$ and $\hat{?} \equiv (-\inf, \inf)$). The operators of SR1 extend the standard operators of the reals (+,−,× and /) to this larger domain, resulting, for example, in the combination of a real and sign algebra. As usual $[r]$ maps a real r to its sign. In SR1 an inequality, such as $C_a > ah$ is expressed by the hybrid equation $[C_a - ah] = \hat{+}$.

Recall that a prominent feature of an equation is a partial order between absolute values of any two terms. Extraction is performed using *Minima*'s algebraic simplification procedures for SR1, substitution of equals or supersets, and for more complicated examples the SR1 hybrid resolution rule, as described in [0]. For example, from the (quantitative) mass balance equation $C_a = ah + a^-$ (I2) and a^- being positive (the qualitative equation $[a^-] = \hat{+}$ (P1)), Minima derives the hybrid equation $[C_a - ah] = \hat{+}$, equivalently $C_a > ah$:

S1)	$C_a - ah = a^-$	Cancellation on I2.
S2)	$a^- \subset [a^-]$	Definition of [ ].
S3)	$C_a - ah \subset [a^-]$	Subst a^- in S1 using S2.
S4)	$C_a - ah \subset \hat{+}$	Subst $[a^-]$ in S3 using P1.
S5)	$[C_a - ah] = \hat{+}$	SR1 simplification of S4.

It follows then from S5 and $[ah] = \hat{+}$ that $|C_a| > |ah|$. Exagerating this feature according to

$$|a| > |b| \rightsquigarrow a \gg b$$

Initial Eqns		Sign Eqns	prominent feature	caricatural assumptions
$h^+ = oh^- + a^-$	(I1)	$[a^-] = \hat{+}, [oh^-] = \hat{+}$	$\lvert h^+\rvert > \lvert oh^-\rvert$	$h^+ \gg oh^-$ (A4)
$h^+ = oh^- + a^-$	(I1)	$[oh^-] = \hat{+}, [a^-] = \hat{+}$	$\lvert h^+\rvert > \lvert a^-\rvert$	$h^+ \gg a^-$ (A3)
$C_a = ah + a^-$	(I2)	$[ah] = \hat{+}, [a^-] = \hat{+}$	$\lvert C_a\rvert > \lvert a^-\rvert$	$C_a \gg a^-$ (A1)
$C_a = ah + a^-$	(I2)	$[a^-] = \hat{+}, [ah] = \hat{+}$	$\lvert C_a\rvert > \lvert ah\rvert$	$C_a \gg ah$ (A2)
$K_w = h^+oh^-$	(I3)		none	none
$K_a ah = h^+a^-$	(I4)		none	none

produces $C_a \gg ah$, which is equivalent to assumption A2 of section . The derivation of each feature and its corresponding caricature is summarized in the above table.

Deriving the caricature of a regime

Given a set of caricatural assumptions defining a subregime, the order of magnitude algebra system *Estimates*[0] is used to derive the caricature of the initial equations, the simplified concentration equations and the domain of validity. First, the assumptions are used to simplify the initial equations, resulting in a set of caricatural equations. Caricatural assumptions and equations are expressed in Estimate's order of magnitude algebra. The types of dominance relations used earlier, $a \ll b$ and $a \approx b$, are captured as algebraic equations in Estimates: $a \gg b \equiv b \subset \epsilon a$, and $a \approx b \equiv a \subset (1+\epsilon)b$, where ϵ denotes a set of (positive and negative) infinitesimal values.[3]

Given a set of caricatural assumptions, Estimates produces a set of caricatural equations, by applying the assumptions to each initial equation using order of magnitude simplification, substitution of superset and qualitative resolution[0]. For example, consider the pair of caricatural assumptions: $C_a \gg ah$ (A2) and $h^+ \gg oh^-(A4)$, which correspond to the example at the beginning of the paper. Applying A2 to the mass balance equation $C_a = ah + a^-$ (I2), Estimates derives the caricatural equation $C_a \approx a^-$ (I2') through the following sequence:

T1)	$ah \subset \epsilon C_a$	Estimates equation for A2.
T2)	$a^- \subset -\epsilon C_a + C_a$	Subst ah in I2 with T1.
T3)	$a^- \subset (1+\epsilon)C_a$	Simplification of T2.
T4)	$a^- \approx C_a$	Relation equivalent to T3.

Likewise, applying A4 to charge balance $h^+ = a^- + oh^-$ (I1) results in $h^+ \approx a^-$ (I1'). Applying these assumptions to I3 and I4 provides no simplification.

Next the concentration equations are derived. Given an equilibrium concentration, such as h^+, and the caricatural equations just derived, Estimates is used to solve for h^+ in terms of the givens K_w, K_a and C_a, producing $h^+ \approx C_a$ (E1):

U1)	$(1+\epsilon)h^+ - a^- \supset 0$	Estimates eqn for I1'.
U2)	$(1+\epsilon)C_a - a^- \supset 0$	Estimates eqn for I2'.
U3)	$(1+\epsilon)h^+ - (1+\epsilon) \supset 0$	Resolve a^- in U1,U2.
U4)	$h^+ \approx C_a$	Reln equiv to U3.

Equilbrium concentrations for ah, a^- and oh^- are derived analogously.

Finally, each bound of the domain of validity corresponds to one of the caricatural assumptions, and is derived using Estimates through a process similar to the above. A boundary is derived from an assumption using the caricatural equations (through qualitative resolution) to eliminate the equilibrium concentrations, resulting in a constraint between givens (K_a,K_w and C_a).

For example, from $C_a \gg ah$ (A2) Estimates derives $K_a \gg C_a$, using I4 ($K_a ah = h^+a^-$), I2' ($h^+ \approx C_a$), and E1 ($C_a \approx a^-$):

S1)	$C_a \gg h^+a^-/K_a$	Resolve ah in A2,I4.
S2)	$C_a \gg h^+C_a/K_a$	Resolve a^- in S1,I2'.
S3)	$C_a \gg C_a^2/K_a$	Resolve h^+ in S2,E1.
S4)	$K_a \gg C_a$	Simplify S3.

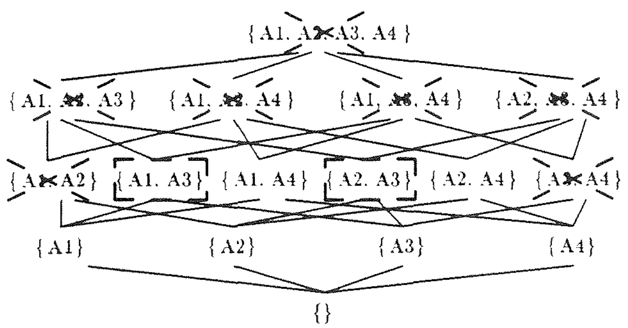
The bounds and concentration equations derived through these processes correspond exactly to those in the example of section .

As a final note, in some cases a set of caricatural assumptions will be mutually inconsistent, for example, as we pointed out earlier for $\{A1, A2\}$. This is recognized when one of the caricatural equations derived by Estimates is not self consistent — for example, from $\{A1, A2\}$ Estimates derives $C_a \gg C_a$ — or is inconsistent with the inequalities derived by Minima.

Creating the patchwork

The relation between caricatures of different regimes has a variety of interesting properties, some having important computational consequences for caricatural modeling. Interrelationships between sets of assumptions can be visualized using a subset/superset lattice:

[3] Intuitively ϵa denotes the set of all values much smaller than a, and $(1+\epsilon)a$ denotes all values close to a.



First, note that at the bottom the lattice is rooted in the original model — since there are no assumptions no exaggeration has been performed. And as we move upwards through the lattice the models become simpler, since each assumption makes an additional term insignificant, which then drops out of the equations.

Second, although models higher in the lattice are simpler, their domain of validity is more restrictive. Since each caricatural assumption introduces a subregime boundary, the region corresponding to the domain of validity of one caricature is a subset of those for any caricature appearing below it in the lattice.

Third, when moving up the lattice additional assumptions do not always result in simplification. For example, $\{A2, A3\}$ produces the same equation for h^+ as does $\{A3\}$.[4] This explains why Schaum's outline includes a region, $R3$, for $\{A2\}$ but no region for $\{A2, A3\}$ (see the region diagram of section). The same argument applies to the absence of $\{A1, A3\}$. These eliminated sets are depicted by squares in the lattice. Likewise, additional assumptions do not always restrict the domain of validity, in particular when the boundary they introduce is outside the existing region.

Finally, while all caricatures could be generated by simply repeating the approach of the previous section on all combinations of caricatural assumptions, the different combinations share two properties that can be exploited to make this process more efficient. First, by monotonicity each superset of an inconsistent set of assumptions is also inconsistent. Thus to avoid exploring potentially large sections of the lattice, we create caricatures starting at the bottom of the lattice and move monotonically upwards, ignoring anything above an inconsistent set. In our example, of 16 potential sets of assumptions, 9 prove consistent, 2 are explicitly demonstrated inconsistent, and 5 are supersets of these, and thus need not be explored. The 7 inconsistent sets are marked by X's in the lattice. Finally, caricatures can also be generated incrementally by exploiting monotonicity. Given the caricature C for a set of assumptions S (in particular C contains the caricatural assumptions and caricatures of initial equations), the caricature of its immediate supersets, $S \cup \{A\}$ are computed by further exagerating C using assumption A.

[4]But this depends on how many of the equilibrium concentrations we are interested in. $\{A2, A3\}$ may allow additional simplification over $\{A3\}$ for other species.

Related Work

An important distinguishing feature of our work, here and in [0] is the emphasis on model *generation*. In constrast to the perspective here on modeling as an emergent process and the focus on quantitative descriptions of dominant behaviors, critical abstraction [0] focuses on extracting qualitative features of models sufficient to understand a behavior of interest. Our ultimate goal is a generative modelling approach that bridges these extremes.

There is a large body of complementary work on the problem of selecting between existing models, which can exploit the models that generative modelling creates. [0, 0] use a graph of model relationships and sensitivities to help select among models. [0, 0, 0, 0] select appropriate models for the constituents of a device, exploiting information supplied about simplifying assumptions, the quantities being observed, and in the last two cases the desired accuracy of the approximations.

Additionally, Caricatural modeling demonstrates the power of qualitative algebraic skills, in particular the use of Hybrid qualitative/quantitative algebra to reason about critical features, and order of magnitude algebra to reasoning about dominance. Several other techniques may be applicable to these two subtasks [0, 0] and [0, 0].

Third, [0] explores the idea of partitioning state space and approximating behavior through a set of piecewise linear approximations. A concern here is that linearization throws away some features of behavior that are particular important, such as the geometric coupling that results from the product of two variables. By concentrating instead on dominant behaviors, caricatural modeling avoids this limitation while still achieving simplicity. The idea of exageration has also been applied to simplifying the DQ analysis problem, as explored by [0, 0].

Finally, two pieces of research on acid-base chemistry are relevant: [0] uses Estimate's order of magnitude algebra to simplify equilibrium equations, but does not capture what we find most interesting, the generation of the caricatural assumptions and the patchwork of regimes. In contrast to our focus on equilbrium behavior, the kineticist's workbench [0] extracts features of the dynamics of chemical reactions.

Acknowledgments

We would like to thank Johan de Kleer for feedback during our hour of need.

References

H. et al. Abelson. Intelligence in Scientific Computing. *Comm. ACM*, 32:546–562, May 1989.

D. Beckwith, editor. *Analytical Chemistry*. McGraw-Hill, NY, 1985.

Dan Weld. Automated Model Switching: Discrepancy Driven Selection of Approximation Reformulations. Technical Report 89-08-01, Department of Computer Science and Engineering, University of Washington, 1989.

Dan Weld. Reasoning about Model Accuracy. Technical Report 91-05-02, Department of Computer Science and Engineering, University of Washington, 1991.

E. Davis. Order of magnitude reasoning in qualitative differential equations. TR 312, NYU Computer Science Department, 1987.

B. Falkenhainer and K. Forbus. Compositional modeling: finding the right model for the job. *Artif. Intell.*, 51, 1991.

M. L. Mavrovouniotis and G. Stephanopoulos. Formal order-of-magnitude reasoning in process engineering. *Computer Chemical Engineering*, 12:867–880, 1988.

D. McAllester. Algebraic approximation. In *Proceedings IJCAI-81*, pages 624–628, 1981.

P. Nayak. Validating Approximate Equilibrium Models. In *AAAI Model-based Reasoning Workshop*, LA, CA, 1991.

P. Nayak, L. Joskowicz, and S. Addanki. Automated Model Selection Using Context Dependent Behaviors. In *International Workshop on Qualitative Reasoning*, Austin, TX, 1991.

O. Raiman. Order of magnitude reasoning. *Artif. Intell.*, 51, 1991.

E. Sacks. 'Hierarchical reasoning about inequalities. In *Proceedings AAAI-87*, pages 649–654, Seattle, WA, 1987.

E. Sacks. 'Piecewise Linear Reasoning. In *Proceedings AAAI-87*, pages 655–659, Seattle, WA, 1987.

Sanjaya Addanki and R. Cremonini and J. S. Penberthy. Reasoning about Assumptions in Graphs of Models. In *IJCAI-89*, Detroit, MI, August 1989. Morgan Kaufmann.

Scott Bennett. Approximation in Mathematical Domains. In *IJCAI-87*, pages 239–241, Milan, Italy, August 1987. Morgan Kaufmann.

R. Simmons. 'Commonsense' arithmetic reasoning. In *Proceedings AAAI-86*, pages 118–124, Philadelphia, PA, 1986.

D. S. Weld. Exaggeration. In *Proceedings AAAI-88*, pages 291–296, St. Paul, MN, 1988.

B. C. Williams. Capturing how things work: constructing critical abstractions of local interactions. In *AAAI Workshop on the Automatic Generation of Abstractions and Approximations*, Boston, MA, 1990.

B. C. Williams. Interaction-based Invention: Designing Novel Devices from First Principles. In *AAAI*, July 1990.

B. C. Williams. A theory of interactions: unifying qualitative and quantitative algebraic rasoning. *Artif. Intell.*, 51, 1991.